Essays in Public and Labor Economics

by

Teodora Nikolaeva Tsankova

Thesis Submitted

for the Degree of Doctor of Philosophy in Economics

to the University of Warwick, Department of Economics

July 2020

Contents

List of Figures

List of Tables

Acknowledgements

I would like to take this opportunity to express my gratitude to everyone who has helped me throughout my years as a doctoral student.

I would first like to thank my supervisors for their continuous support during the past four years. I am very grateful to Clément Imbert for generously offering me his time and attention whenever I needed it. His feedback has been invaluable to me as a Master's and PhD student. I am very grateful to Sharun Mukand for offering me guidance and honest advice throughout.

I also thank all members of the Economics Department at the University of Warwick who have helped me during my years at the University of Warwick. I am in particular grateful for the academic support I received during my last year by James Fenske and Manuel Bagues. I acknowledge the help of Natalie Deven and Maryanne Heafey and all members of the Centre for Competitive Advantage in the Global Economy. I am thankful for the support and friendship of my course mates Andrea, Andrei, Anthony, Carlo, Daniel, Haseeb, Horng and Shi.

Finally, I thank my family. I am particularly grateful to Iliyana who has helped me in every possible way for the last six years. I am truly lucky to have her as my sister. I am thankful to my parents, Miglena and Nikolay, for continuously supporting me during my studies at all levels and to Vlatko for his patience.

Declaration

This thesis is submitted to the University of Warwick in accordance with the requirements of the degree of Doctor of Philosophy in Economics. I declare that it has not been submitted for a degree at another university. Chapter 1 and Chapter 2 are co-authored with Mirjam Bächli (University of St.Gallen). Chapter 3 is joint work with Jan-Emmanuel De Neve (University of Oxford), Clément Imbert (University of Warwick), Johannes Spinnewijn (London School of Economics) and Maarten Luts (FPS Finance).

July 2020

Abstract

This thesis explores topics in public and labor economics summarized below.

Chapter 1 explores drivers of native support for immigration. It studies the importance of labor market concerns and the role of labor protection in shaping native preferences over migration policies. We look at the Swiss context and national votes which took place in the period from 2000 to 2014. Our results show that a higher migrant exposure reduces pro-immigrant vote shares in municipalities with a relatively low-skilled native population. The negative response is mitigated under higher levels of labor market protection as measured by collective bargaining coverage. Consistent with labor market concerns driving support for immigration, we find that immigration reduces wages of low-skilled natives, but this effect is weaker under higher collective bargaining coverage.

Chapter 2 investigates how the introduction of free movement of workers affects enrolment of natives in tertiary education. We exploit a Swiss policy change that led to a significant increase in the share of cross-border commuters in local employment in border regions of Switzerland. Our results show a rise in enrolment in affected relative to non-affected regions in the post-reform period driven by universities of applied sciences. Consistent with returns to education driving enrolment decisions, we observe a decrease in wages to upper-secondary degrees and a rise in wages for tertiary educated workers. Furthermore, we link occupations to study fields and divide subjects according to how much they are affected by the inflow of commuters. Our results show that enrolment in less affected fields rises. These are non-STEM subjects which build skills that are less likely to be transferable across borders.

Chapter 3 studies the impact of simplification, deterrence and tax morale on tax compliance. We ran five natural field experiments varying the communication of the tax administration with the universe of income taxpayers in Belgium throughout the tax process. A consistent picture emerges across experiments: (i) simplifying communication substantially increases compliance, (ii) deterrence messages have an additional positive effect, (iii) invoking tax morale is not effective, and often backfires. A discontinuity in enforcement intensity, combined with the experimental variation, allows us to compare simplification with standard enforcement measures. We find that simplification is far more cost-effective, allowing for substantial savings on enforcement costs.

1 Does Labor Protection Increase Support for Immigration? Evidence from Switzerland

with Mirjam Bächli

1.1 Introduction

The number of international migrants has risen by nearly seventy percent since 1990, reaching 272 million people globally (UNPD, 2019). A small number of countries, mostly high-income, have received a disproportionate share of immigrants. Among OECD members the foreign population makes up approximately nine percent of the population (OECD, 2020). At the same time, immigration has come to the center of political debates in a number of these countries. Anti-immigrant rhetoric dominated recent elections in the United States and several European countries, and the debates leading up to the Brexit referendum. The wide voter support such campaigns receive is evidence of a rising concern about how foreigners are integrated into society and the labor market of the receiving country.

We investigate the role of labor market concerns in shaping preferences over migration policies. Our main contribution is to test whether protecting native working conditions affects support for immigration. We focus on Collective Bargaining Agreements (CBAs) which set wage and working conditions for the contracting parties. Our results show that collective bargaining coverage determines how native support for immigration and native labor market outcomes respond to the local presence of migrants.

Switzerland offers a favorable setting to study our research question. With 24 percent foreigners in the population, the country ranks second among the OECD member states (OECD, 2020). Swiss direct democracy gives voters a say on national policies. We focus on proposed changes to migration policy with implications for aggregate number of foreign residents. The votes took place between 2000 and 2014. Vote outcomes reveal the degree of native support for immigration. Recent changes to migration policies have been implemented conditional on accompanying measures, including the enforcement of existing collective bargaining agreements. The country does not have a national minimum wage, but relies heavily on collective agreements to regulate industry- and region-specific wage conditions with a coverage rate of 43% in 2016.

The empirical analysis links native labor market outcomes and voting behavior to local exposure to immigrants under different levels of collective bargaining coverage. We use information on generally valid CBAs and employment by industry to build a regional measure of the share of workers employed in an industry with a collective agreement. A number of steps address concerns that agreements are a response to deteriorating working conditions or to the inflow of immigrants. We construct a leave-one-out measure of coverage at the industry level, where we ignore employed in the region, and fix employment by industry at baseline. Similarly, immigrants could self-select into regions with more positive attitudes towards immigrants or better labor market conditions. To mitigate these concerns, we rely on an instrumental variable strategy that uses past settlement patterns to allocate migrants to regions within Switzerland.

We start by comparing native skill levels to those of immigrant workers, and calculate collective bargaining coverage rates by level of skill. Immigrants are overrepresented to the left of the skill distribution and underrepresented to the right. Similarly, low-skilled natives are more likely to be employed in industries with a CBA than high-skilled natives. Collective agreements, therefore, protect labor market outcomes for the subset of natives who are likely to compete against foreign workers.

The analysis of voting outcomes reveals a positive but insignificant effect of a higher immigrant exposure on the share of pro-immigrant votes. Native educational attainment determines how vote outcomes respond to the presence of migrants. Specifically, as the skill level of the native population declines, the response becomes more negative. At the low end of the skill distribution we estimate that a rise in immigration equal to 1 percent of the native population leads to a decline in pro-immigration vote shares of 0.2 percentage points. Results show that in these municipalities the negative response is weaker under higher collective bargaining coverage. At the upper end of the skill distribution the response to a higher local presence of immigrants is positive and does not depend on the level of labor market protection.

To assess the relevance of labor market concerns as a determinant of voting behavior, we turn to native labor market outcomes. Results show that a rise in immigration is linked to a reduction in wages for low- to medium-skilled workers. For natives at the lower end of the skill distribution a rise in the number of migrants equal to 1 percent of the native population decreases wages by 0.6 to 0.7 percent. Results confirm that collective bargaining agreements mitigate these negative effects – at high levels of coverage estimates are around -0.4 to

-0.5 percent. Similarly, we find that negative employment effects are mitigated by a higher CBA coverage.

Our main contribution is to link vote outcomes to the presence of immigrants and the underlying labor market institutions. To the best of our knowledge, ours is the first attempt to draw a causal link between labor protection and preferences over migration policies. The analysis of wage and employment outcomes complements the voting results and provides evidence that labor concerns affect support for immigration.

The determinants of native attitudes towards immigration have been studied using social survey data with mixed evidence. Exposure to migrants could theoretically reduce prejudice as suggested by intergroup contact theory (Allport et al., 1954). Schindler and Westcott (2017) find that the historic presence of black American military units in the United Kingdom during the Second World War is linked to lower stated prejudice and lower implicit bias towards blacks. In contrast, Dustmann and Preston (2001) find that racial intolerance can explain negative attitudes towards ethnic minorities again in the UK. Card et al. (2012) argue that the desire to preserve a country's customs and culture is the main driver of anti-migrant sentiments. Similarly, using Swiss data Hainmueller and Hangartner (2013) and Diehl et al. (2018) find that preferences over migration policies vary with the country of origin of the migrant population considered.

Exposure to migrants could also increase anti-migrant sentiments if it raises economic concerns among the native population. A number of studies link concerns over the fiscal burden of immigration to negative attitudes towards migrants (see Dustmann and Preston, 2007; Facchini and Mayda, 2009; Alesina et al., 2018). Another strand of the literature argues that natives who are likely to compete against foreigners in the labor market hold more negative attitudes (see Scheve and Slaughter, 2001; Mayda, 2006; O'Rourke and Sinnott, 2006; Ortega and Polavieja, 2012). Pecoraro and Ruedin (2019) argue that occupation-level exposure to foreigners affects negatively native attitudes particularly when job prospects are worse in Switzerland. In contrast, d'Hombres and Nunziata (2016), Cavaille and Marshall (2019) and Margaryan et al. (ming) find that education decreases the probability of holding anti-migrant views but this is not driven by a labor market channel. While we also proxy exposure to foreign competition in the labor market with skill level, we use regional variation in labor protection to credibly test the relevance of labor market concerns.

A growing literature links election outcomes and exposure to immigrants. Evidence from Austria (Halla et al., 2017), Denmark (Dustmann et al., 2019), Ger-

many (Otto and Steinhardt, 2014), Italy (Barone et al., 2016), France (Edo et al., 2019) and United States (Tabellini, 2019) suggests that higher local migrant presence is associated with more votes for right-wing parties. Similarly, Cavaille and Ferwerda (2017) argue that support for the far-right rose in response to a regulatory change that granted non-EU migrants access to public housing in Austria. Guiso et al. (2017) link votes for populist parties to economic insecurity induced by immigration. On the other hand, Mayda et al. (2018) find that Republican vote shares are lower in US counties with higher share of migrants but the response to low-skilled immigration is positive. Steinmayr (ming) finds that long-term interaction reduces far right vote shares in the context of Austria, evidence in line with the contact hypothesis, while short interaction has the opposite effect.

Rather than looking at stated attitudes towards migrants or inferring support for immigration from election outcomes, we focus on popular votes on immigration regulation. Facchini and Steinhardt (2011) argue that votes on immigration policy in the US House of Representatives are driven by labor market concerns as proxied by the skill level of the constituency. The Swiss direct democracy allows us to observe directly voting behavior of the native population. Most relevant to our work is the study of Brunner and Kuhn (2018) who look at Swiss votes related to immigration regulation. Their results point at a sizeable increase in anti-immigration vote shares as a response to the presence of culturally different migrants in the municipality. Our paper focuses on labor market concerns and asks whether protective labor market institutions are effective in alleviating them.

The literature has found mixed evidence of how immigration affects native wages (see Borjas, 2003; Ottaviano and Peri, 2012a; Dustmann et al., 2016a). Using a skill-cell approach and Swiss data, Gerfin and Kaiser (2010) document positive effects for the low-skilled and negative for the high-skilled natives, while Basten and Siegenthaler (2019) find no significant wage effects. Following Dustmann et al. (2016a), we do not allocate migrants to skill groups but take a geographic area approach. Our empirical strategy is similar to Beerli et al. (2018) who look at exposure to high-skilled cross-border commuters in Switzerland and find a positive effect on the wages of tertiary educated natives. In contrast, we study resident foreigners.

Union membership received significant attention in the economics literature in the 1990s (see Card, 1996; Lemieux, 1998). With respect to collective bargaining agreements, recent papers have found mixed evidence on the wage effects (Card and De La Rica, 2006; Gürtzgen, 2016) and some evidence of negative employment effects (Kahn, 2000; Magruder, 2012). Our focus is on how collective

bargaining affects native labor market outcomes in the context of rising immigration. In an early paper focusing on European states, Angrist and Kugler (2003) argue that labor regulation can protect some native workers from immigrant competition, but it can also lead to worse employment outcomes. Recent work investigates employment protection (D'Amuri and Peri, 2014), fixed versus indefinite term contracts (Edo, 2016) and minimum wages (Edo and Rapoport, 2019). A number of papers focus on negative employment effects of immigration under rigid wages (see Boeri and Brücker, 2005; Brücker and Jahn, 2011; Brücker et al., 2014). In a meta analysis Foged et al. (2019) argue that institutional differences are vital in reconciling findings from different countries. Collective bargaining, specifically, is not found to have a significant effect. In contrast, we use within country variation in labor market protection and take steps to alleviate endogeneity concerns with respect to collective bargaining coverage.

Recent work on labor market effects of immigration is often motivated by anti-migrant sentiments. We contribute to this field by examining labor market effects and explicitly looking at support for immigration. This approach is also followed by Tabellini (2019) who studies the period of Mass Migration in the United States and argues that not economic concerns but cultural differences are the drivers of hostility against migrants in that context. The combination of a voting and labor market analysis, and the variation in labor market protection allow us to draw a link between labor market concerns and preferences over migration policies.

The remainder of the article is organized as follows: in Section 1.2 we discuss context-specific immigration policies and collective bargaining agreements, and data used; in Section 1.3 we describe the empirical strategy and present basic trends in the data; in Sections 1.4 and 1.5 we present our results; Section 2.6 concludes.

1.2 Context and Data

1.2.1 Swiss Context

Migration regulation Swiss migration policies differentiate migrants on the basis of country of origin. Current policies in place imply that European Union and European Free Trade Association countries face preferential treatment relative to third-country nationals.

The Agreement on the Free Movement of Persons (AFMP) was negotiated as a part of the Bilateral Agreements I with the European Union. Initially it applied

to workers from EU-15/EFTA member states and was later extended to new EU members.[1] The policies were implemented specifically for country groups and over several years. EU-15/EFTA and EU-10 members have enjoyed unconditional free movement of persons since 2014, EU-2 members since 2019. In contrast, immigration policy applicable for non-EU/EFTA workers is strictly regulated. Rules are guided by the Federal Act on Foreign Nationals and Integration which came into force in January 2008. Quotas are decided on an annual basis by the federal government.[2] Prerequisites for a working permit to be granted include a high skill level, non-violation of the local priority requirement, wage and working conditions that correspond to local and professional standards to prevent wage dumping.[3] Note that different rules apply to foreign nationals who migrate for family reunion reasons and to asylum seekers.

The Swiss direct democracy allows its citizens over eighteen years of age to take part in compulsory and optional referendums, as well as popular initiatives.[4] Voters can challenge newly approved migration policies by the parliament with optional referenda and in addition propose changes through popular initiatives. Since only Swiss nationals are eligible to vote, results reflect native preferences. Popular votes are scheduled three to four times per year and each eligible voter receives a voting booklet with details of the proposal. Hence, voters understand well the principles of direct democracy and have access to all relevant information to make an informed choice.

According to data from the Swiss State Secretariat of Migration (SEM), the number of foreign residents in Switzerland rose by more than 53% between 2000

[1]EU-15 member states are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, United Kingdom; EFTA are Iceland, Liechtenstein and Norway; EU-10 are Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia and Slovenia; EU-2 are Bulgaria and Romania; EU-28 includes EU-15, EU-10, EU-2 and Croatia.

[2]Determinants are economic demand, utilization of the previous year's quotas, hearings of the cantonal authorities and employer/employee representatives.

[3]High-skilled people are defined as tertiary educated with several years of professional experience as a manager or a specialist. To fulfil the local priority requirement employers need to present proof that there are no other suitable Swiss or EU/EFTA candidates available for the specific position.

[4]Constitutional amendments or accessions to supranational organizations are by default subject to a compulsory referendum. Optional referendums can challenge an act passed by parliament. A popular majority is sufficient for approval. Popular initiatives allow voters to submit proposals that will be incorporated into the federal constitution conditional on being accepted. A sufficient condition for a popular initiative is that 100,000 signatures are collected within 18 months after having fulfilled some formalities that are confirmed by the Federal Chancellery. For comparison, a minimum of 50,000 signatures have to be collected within 100 days after the official publication of the act for an optional referendum to be called. Alternatively, a minimum of eight cantons can demand a vote. A majority of voters and a majority of cantons must vote in favor of the initiative for it to be approved. A double majority is also required for a compulsory referendum to pass.

and 2018 to above 25% of the population. This observed rise in immigration is largely driven by foreigners coming from EU/EFTA countries. They account for 69% of all immigrants at the end of the period. There are significant differences in motives for migration among the EU/EFTA and third-country nationals. In 2018 64.7% of EU-28/EFTA nationals entered for employment reasons and only 22.8% for family reasons. In contrast, 10% of the inflow of non-EU/EFTA nationals in the same year came for reasons of employment, while 47.3% of them entered for reasons of family reunification. This can be linked to the policies in place which make it difficult for non-EU/EFTA nationals to acquire a working permit.

Collective bargaining The Swiss labor market is considered relatively unregulated – it ranks 35 out of 42 countries in 2013 according to the employment protection legislation index of the OECD.[5] Collective bargaining agreements (CBA) are a common tool to set working conditions in Switzerland. A collective bargaining agreement is a fixed-term contract with normative provisions such as minimum wages, working hours, holidays and wage eligibility during sickness, motherhood and military service. These conditions are binding for the contracting parties, which are the involved employers and employees. Unions often negotiate these conditions on behalf of the employees. However, an employee must not necessarily be union member to benefit from the negotiated clauses because of firms applying the conditions to the entire workforce or of extension mechanisms making the agreement generally valid for an occupation or industry.

According to the Survey on Collective Labour Agreements (SCLA), the number of covered employees rose from 1.26 million to 2 million between 1999 and 2014. In the end of the period around 41% of the employed are covered.[6] The upward trend is driven by a rising number of workers covered by agreements with minimum wage clauses, which account for 87% of the covered population in the end of the period. Agreements often specify minimum wage levels for individuals by educational attainment, while some also condition on experience. The rise is due to new CBAs being signed but also existing CBAs made generally valid.

Since 1956 it is possible to declare an existing CBA generally valid.The procedure starts with a written request from the contracting parties. Certain conditions must be met for a request to be approved – importantly there is a quorum requirement where initial contracting parties should account for a majority of

[5]The index measures the strictness of regulation on individual and collective dismissals. The US is ranked most liberal.

[6]While the number of covered includes employers as well as employees, we consider the former as insignificant.

covered employers and employees.[7] Once declared generally valid, agreements are binding for everyone in the industry or the occupation. Although only a small share of all CBAs are generally valid, they account for 50% of all covered workers. The total number of such agreements rose from 36 in 2000 to 71 in 2014, and the number of total workers covered doubled. We concentrate on generally valid CBAs as we think of such agreements as setting industry-wide standards. Such agreements are divided into cantonal and national ones, where at least two cantons are involved in the latter. Minimum wages as a form of labor protection can alleviate concerns over labor market outcomes, so we focus on agreements with minimum wage clauses (86% of all generally valid agreements).

Labor market considerations played a dominant role in the discussions prior to the votes. The AFMP, specifically, was predicted to significantly increase the inflow of EU/EFTA workers with potentially negative consequences for native workers. Implementation was, therefore, conditional on accompanying measures aimed at ensuring that all firms and employees adhere to the Swiss wage and working conditions. The accompanying measures came into force in 2004, and include systematic wage controls to prevent abusive wage undercutting and sanctions for breaching the rules. The systematic wage controls are conducted by the parity and tripartite commissions. It is the responsibility of the parity commission (including representatives from unions and employer associations) to check whether CBA conditions are fulfilled and hence strengthen the enforcement of agreements. If violations are found, an existing CBA can be declared generally valid in a simplified procedure[8] or if it already exists, fines and sanctions are imposed. The cantonal tripartite commissions consist of representatives from unions, employer associations, and the canton itself. They conduct risk-based wage and working condition controls in industries without a CBA. Parental union organizations recommended a vote in favor of the AFMP as accompanying measures were deemed sufficient to protect working conditions.

[7] The following conditions defined by law must be fulfilled for an agreement to be declared generally valid: (1) necessity; (2) non-infringement of general interest and minority interests considered; (3) quorum conditions – more than half of the employers being covered by the generally valid CBA must be part of the current CBA; more than half of the employees being covered by the generally valid CBA must be part of the current CBA; the employers involved in the current CBA must employ more than half of the employees that will be covered under the generally valid CBA.

[8] In a simplified procedure only the third quorum condition has to be satisfied. Note that the simplified procedure is rarely used – in the period 2004–2018 it was applied only in four cases.

1.2.2 Data

We use a combination of administrative data and large-scale surveys.

Voting outcomes We look at votes that relate to Switzerland's migration policy towards EU/EFTA nationals or third-country nationals. The six votes we focus on have direct implications for immigration levels and took place in the period 2000–2014. Specifically, we include four optional referenda, which relate to the AFMP and the Federal Act on Foreign Nationals and Integration (AuG), and two popular initiatives, which propose quantitative restrictions (for a list of votes see Table A1).

Results of national votes at the municipality level are provided by the Federal Statistical Office (FSO).[9] Available information includes the number of citizens with the right to vote, participation and acceptance rate. In line with Brunner and Kuhn (2018), we classify proposals as pro- or anti-immigration based on implications for aggregate immigration levels. The outcome of interest is the share of pro-immigration votes in a municipality and we also look at participation rates.

We supplement these data with information from the Vox Survey (Vox Survey, 2019). This is a post-vote telephone survey covering eligible voters. While it has been conducted since 1981, we restrict the sample to the six votes used in our main analysis. The questionnaire asks whether and how respondents voted in a specific vote, about demographic characteristics, income level and a set of attitudinal questions. We link self-reported voting behavior to stated attitudes towards foreigners in the country.

Labor market outcomes The Swiss Earnings Structure Survey (SESS) is a large-scale firm survey conducted biennially in the month of October between 1994 and 2016. It is a repeated cross-section of firms covering the secondary and tertiary sectors of the economy. Respondents provide information about a random subset of employees. The number of workers covered depends on firm size, with information available for at least one third of all workers. In 2014 the survey included about 32,000 public and private enterprises with approximately 1.6 million workers. At a firm level there is information about the region where the firm is located, industry and size. The SESS has information on the gross

[9] Out of the 2,222 municipalities in 2018, seven do not have an own voting office leaving us with a sample of 2,215 municipalities.

hourly wages of individual workers and their educational attainment.[10] We differentiate between three skill levels based on highest education attained – at most up to lower-secondary, upper-secondary and tertiary education. Data allows us to distinguish native from foreign workers, and within the latter group foreigners with different permit types. The sample is limited to employees 18–65 years of age, working in private sector enterprises with available region of work and permit type as well as gender. We collapse the employee-level data at the regional level using survey weights provided. Our main outcome of interest is the mean hourly wage by skill level where we proxy skill with educational attainment and percentiles of the native skill distribution (see Dustmann et al., 2013).[11]

While the SESS covers only employed individuals, the Swiss Labor Force Survey (SLFS) includes individuals aged 15 years and older. The survey was conducted annually in the second quarter of the year from 1996 to 2009 and quarterly in the period 2010 to 2018. For consistency, annual data is used over the full period. In 2014 around 125,000 interviews were conducted. Information about municipality of residence, demographic characteristics (sex, age and marital status), educational attainment and employment outcomes is available for the household head. We limit the sample to individuals in the age group 18–65. Employment is defined as being employed for a salary, by a family member or self-employed. The main outcome of interest, constructed using survey weights, is the native employment rate in a region – the number of employed relative to population 18–65 years of age. We construct outcomes by educational levels defined as in the analysis of wage outcomes.

Immigration Our main data source is the Swiss Central Migration Information System (ZEMIS). Among migrants, we use information on individuals with short-term (L), resident (B) and settled status permit (C).[12] Individuals are covered if they reside in the country on December 31 for the period 1996–2018. The data are provided at the municipality level before 2002 and at the individual level after. They offer information on the flow and stock of migrants by country of citizenship, permit type, gender, age and civil status.

To calculate local exposure to migrants, we combine these data with information

[10]Gross hourly wages include social contributions, Sunday or night work compensation, 1/12 of 13th salary and other non-periodic payments but exclude overtime pay. Real values are constructed using CPI data indexed to December 2015.

[11]Following Dustmann et al. (2013), we trim observations above the 99th and below the 1st percentile of the wage distribution in each region.

[12]Provisionally admitted foreigners (F permit), asylum seekers (N permit) and people in need of protection (S permit) are excluded. These types of permits are granted only to non-EU/EFTA nationals. We also exclude cross-border commuters (G permit) who come mainly from EU/EFTA member states.

on population size at the municipality level from the FSO. Our measure is the number of migrants in a region in a year divided by the native population in 2000. By fixing the denominator we abstract from issues related to native mobility as well as naturalization of foreign nationals. Additionally, we use data on migrant stocks by citizenship at the municipality level from the 1970 and 2000 census and 2000 census data in constructing demographic controls for the vote analysis.

Collective bargaining The State Secretariat of Economic Affairs (SECO) provides a list of the universe of generally valid agreements from June 2000 onwards. We restrict the sample to CBAs with a minimum wage clause, which account for 89% of all agreements. Our database shows for each CBA the name, the period when it was in force, its geographic coverage and the 4-digit NOGA-08 industry. We calculate the share of workers employed in an industry with a CBA by combining time-varying information about coverage at the industry level and time-invariant native employment by industry in a region.

$$Sh\,CBA\,Cov_{r,t} = \sum_{i=1}^{I} Sh\,Empl_{i,r,1995} \times \mathbb{1}\{CBA_{r,i,t} = 1\}$$

A region is indicated with r, t is year, i is industry at the 3-digit NOGA-08 level and I the total number of such industries (259 for 3-digit). The first term on the right-hand side is the share of native employees in region r that work in industry i in 1995. The second term is a dummy variable equal to one if there is a generally valid CBA in region r in industry i and year t.

1.3 Empirical Strategy and Trends

1.3.1 Empirical Strategy

We are interested in how the regional exposure to immigrants affects views on immigration policies and labor market outcomes. Our main contribution is to investigate whether these effects depend on the level of collective bargaining coverage. The empirical analysis builds on the following two regression equations.

$$y_{r,t} = \alpha m_{r,t} + \mathbf{X}'_{r,t}\gamma + \delta_r + \delta_t + \varepsilon_{r,t} \tag{1.1}$$

$$y_{r,t} = \alpha_1 m_{r,t} + \alpha_2 m_{r,t} \times Sh\,CBA\,Cov_{r,t} + \alpha_3 Sh\,CBA\,Cov_{r,t} + \mathbf{X}'_{r,t}\gamma + \delta_r + \delta_t + \varepsilon_{r,t} \tag{1.2}$$

Region r and year t define the unit of observation. In the analysis of voting outcomes r stands for municipality. Municipalities are the smallest administrative units – a total of 2,222. Their large number makes them attractive for the analysis of vote outcomes in the absence of individual-level data. In the labor market analysis the geographical unit is the commuting zone or MS-region.[13] A commuting zone – 106 in total – consists of municipalities that are spatially similar, so obey the principles of small-scale labor market areas.

Outcome variables $y_{r,t}$ measure the share of pro-immigration votes, the natural log of gross hourly native wages and the native employment rate. The latter two outcomes are analysed by educational attainment while wage effects are additionally proxied with percentiles along the native wage distribution. The analysis is separately conducted for five-percentile steps following Dustmann et al. (2013). Our main independent variables are $m_{r,t}$, the migrant exposure measure, and $Sh\,CBA\,Cov_{r,t}$ which measures the extent of CBA coverage.

As control variables we add characteristics of natives that are consistent in the two parts of the empirical analysis to $\mathbf{X}_{r,t}$ – gender, average age and highest educational attainment. In the labor market analysis those refer to time-varying native worker and respondent characteristics from the SESS and SLFS, respectively. In the voting analysis age and education are fixed in 2000 and interacted with a year variable (for an overview of control variables see Table A2). We include region fixed effects, as well as year or referenda fixed effects in the labor market or voting analysis, respectively. Standard errors are clustered at the regional level.

The place of residence of immigrants likely depends on labor market conditions as well as native attitudes towards foreigners and is, therefore, not random. To address this, we follow an instrumental variable approach where we create a shift-share instrument for immigration exposure (see Altonji and Card, 1991; Card, 2001). We fix the share of immigrants from origin o across regions r in 2000 and use levels of overall immigration M by origin o in year t as the yearly shift.[14]

$$\overline{M}_{r,t} = \sum_{o=1}^{O} Sh\,migr_{o,r,2000} \times M_{o,t}$$

[13] MS comes from the French "mobilité spatiale".

[14] We use 162 separate countries of origin in the construction of the instrument. Seventeen countries of origin are not in the 1970 census data, while only one is not in the 2000 census. We allocate migrants from those countries according to initial geographical distribution of migrants with an unknown origin. Alternatively, we can group countries based on geographical location. Results, not reported for brevity, remain consistent.

This identification strategy has been widely applied in the migration literature. The intuition is that past immigration can predict location choice of newly arriving migrants. This strategy addresses the endogeneity problem if past immigration is uncorrelated with current demand shocks. In a robustness check we use the 1970 distribution of migrants by origin which one could argue is less likely to be correlated with current location-specific shifts in demand. Results remain qualitatively similar.

Recent criticisms of the use of shift-share instruments in this context were raised by Jaeger et al. (2018) who argue that estimates are likely positively biased as they reflect dynamic adjustments of economic conditions to previous migration waves. We consider this less of a concern in our context for a number of reasons. First, the origin composition of migrants changed substantially since the 1990s. The correlation between changes in immigrant stock by country from 1990 to 2000 and 2000 to 2010 is 0.11 and not significant.[15] Second, we document negative wage effects for natives in skill groups affected by immigration which is unlikely under dynamic adjustments to past migration.

Other studies raise general concerns with the use of a shift-share type of instrument (see Borusyak et al., 2018; Goldsmith-Pinkham et al., ming). The underlying variation in our instrument comes from 159 origins over a period of fifteen years. Following the literature we compute the Rotemberg weight (RW) for each country. The three origins receiving the highest positive weights turn out to be the same in the analysis of vote, wage and employment outcomes – Germany, France and Portugal. These are cases where estimates tend to be sensitive to misspecification. We exclude these countries in a robustness check when constructing the instrumental variable. Results are shown to be robust.

Generally valid CBAs are attractive for identification purposes as they are binding for firms that did not initiate the agreement, making variation more likely to be exogenous. Although they are publicly available, people are likely not aware of them, especially as individual working contracts can deviate in favor of the worker. Hence, neither native not foreign workers are expected to take generally valid CBAs systematically into account when deciding in which region to live or work. In our baseline measure of coverage we fix the industrial structure to 1995 values. By abstracting from over-time changes in Swiss employment by industry, we rule out variation in industrial structure due to the new CBAs or to the inflow of foreign workers.

[15]The correlation between changes from 1980 to 1990 and 1990 to 2000, on the other hand, is 0.86 indicating that the origin composition of immigrants remained very similar over period. Results are comparable if one looks at changes in the share of migrants by origin group.

In the specification in which we instrument migrant exposure, we use a modified CBA coverage measure to avoid endogeneity issues.

$$\overline{Sh\,CBA\,Cov_{r,t}} = \sum_{i=1}^{I} Sh\,Empl_{i,r,1995} \times Sh\,CBA\,Cov_{i,-r,t}$$

$$where \qquad Sh\,CBA\,Cov_{i,-r,t} = \sum_{r'=1}^{R} Sh\,Empl_{r',i,1995} \times \mathbb{1}\{CBA_{r',i,t} = 1\}, \forall r' \neq r$$

$Sh\,Empl_{r',i,1995}$ is share of industry i's employment in 1995 that is in region r'. $\mathbb{1}\{CBA_{r',i,t} = 1\}$ is an indicator function equal to one if there is a collective agreement in region r' in industry i in year t and zero otherwise. By focusing on individuals employed in other regions, we aim to mitigate concerns about reverse causality or that new agreements are a response to changes in local labor market conditions. The limitation of such a specification is that we can interpret estimates only as reduced form effects.

1.3.2 Summary Statistics and Stylized Facts

Table 1.1 shows summary statistics for the main variables of interest in the empirical analysis.

Voting outcomes as measured by the share of pro-immigration votes are summarized first. There is substantial variation in outcomes across the six votes we consider and some policy proposals faced much higher voter approval than others. For example, the Bilateral Agreements I with the EU which is the first vote we consider (a pro-immigration proposal) was approved by a clear majority of voters. The AuG vote which took place in 2006 (an anti-immigration proposal) also had wide voter support. Additionally, there is considerable variation across municipalities for each of the votes included.

Wage and employment outcomes for native workers at the commuting zone level for the period 2000–2014 are presented next. The mean log gross hourly wage received by native workers is 3.6 (35 CHF in levels). There is a large wage premium to upper-secondary but particularly to tertiary education. For the average region a low-skilled worker earns a gross hourly wage of about 28 CHF while a high-skilled worker earns approximately 48 CHF.[16] The average native

[16]The exchange rate USD/CHF is approximately 1.03 (August, 2019).

Table 1.1: Summary statistics

	N	Mean	Sd	Min	Max
Share pro-immigration votes: 2000 I	2215	0.671	0.119	0.059	0.960
Share pro-immigration votes: 2000 II	2215	0.638	0.091	0.191	0.947
Share pro-immigration votes: 2005	2215	0.558	0.106	0.068	0.824
Share pro-immigration votes: 2006	2215	0.320	0.101	0.063	0.705
Share pro-immigration votes: 2009	2215	0.595	0.110	0.081	0.860
Share pro-immigration votes: 2014	2215	0.495	0.112	0.064	0.810
Mean ln gross hourly wage of natives	848	3.594	0.109	3.246	3.837
...lower-secondary educated	848	3.345	0.082	2.924	3.732
...upper-secondary educated	848	3.526	0.081	3.219	3.729
...tertiary educated	847	3.879	0.103	3.277	4.078
Native employment rate	1590	0.776	0.047	0.332	1.000
...lower-secondary educated	1576	0.451	0.117	0.000	1.000
...upper-secondary educated	1590	0.787	0.059	0.132	1.000
...tertiary educated	1585	0.909	0.053	0.000	1.000
Immigrants to 2000 native population	1590	0.294	0.137	0.038	0.685
Share CBA covered	1590	0.158	0.045	0.074	0.448

Note: The table presents summary statistics for voting and native labor market outcomes, immigrant exposure and collective bargaining agreement coverage. See Table A1 for a description of the votes considered. Voting outcomes are weighed using the number of voters, labor market variables with the number of native workers in 2000 (SESS data) and the number of native respondents 18-65 years of age in 2000 (SLFS data). The migrant exposure measure is weighed with the total population level in 2000 and the share of CBA covered with the number of workers in 2000. SESS, SLFS, migrant exposure and CBA coverage variables are measured at the commuting zone level, vote outcomes at the municipality level. Source: FSO, SECO, SESS, SLFS, ZEMIS.

employment rate is 77.6% and varies widely across skill groups. Among lower-secondary individuals it is 45.1% and among tertiary educated 90.9%.

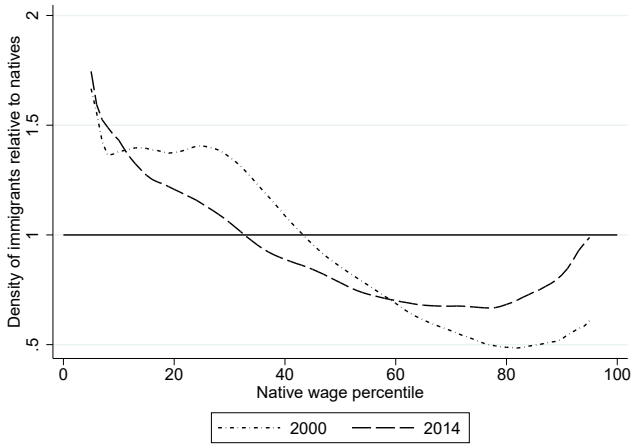
Due to institutional differences, educational degrees are not necessarily comparable across countries. In addition, skills acquired abroad may not be perfectly transferable and, thus, be discounted. Wages allow an alternative view on how education is valued on the labor market. Figure 1.1a follows Dustmann et al. (2013) and plots the share of migrants along the native wage distribution.[17] The horizontal line at 1% is a natural point of comparison as it represents the equal split of natives along own wage distribution. The graph shows that migrants are overrepresented not only at the very bottom of the income distribution but up to the fortieth percentile. Overall, this evidence suggests that natives with a low- to medium-level of skill are the ones who face labor market competition with foreign workers. This is confirmed by Figure A1a which plots the share of native and migrant workers by educational attainment.

The SESS does not provide information on nationality. Hence, we are not able to

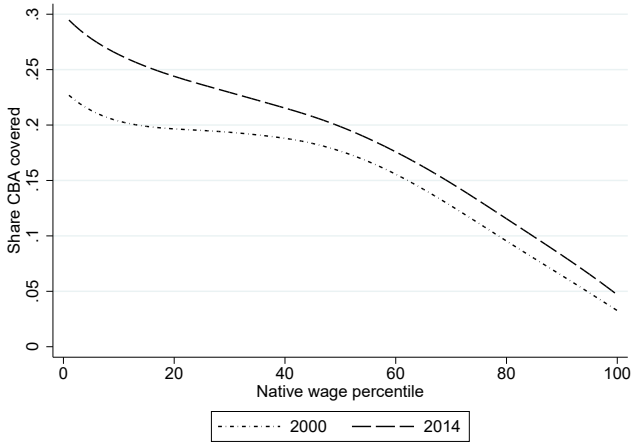
[17]Following Dustmann et al. (2013), the kernel estimates are calculated on the log of the odds of the position in the native distribution and are then transformed.

Figure 1.1: Skill level and CBA coverage

(a) Position of immigrant workers in native wage distribution



(b) Native CBA coverage by position in wage distribution



Note: Figure (a) presents kernel estimates of the density of migrant workers along the native wage distribution. Figure (b) presents a local linear smooth plot of the share of native workers employed in an industry with a CBA by percentile of the native wage distribution. Source: SECO, SESS.

compare the skill of EU/EFTA and non-EU/EFTA to the one of Swiss workers as proxied by educational attainment or wage percentiles. The SLFS database, however, offers information on the occupation of Swiss workers as well as migrants by nationality. In 2014 migrants are overall more likely than Swiss workers to hold occupations at lower levels of the ISCO-08 classification. Therefore migrants tend to hold jobs with lower skill requirements. Furthermore, a comparison between EU/EFTA and non-EU/EFTA nationals shows that this pattern is even more pronounced in the latter group. Therefore, we expect that low-skilled natives are the ones who compete against migrants and particularly against non-EU/EFTA workers.

The share of CBA covered in employed among natives is 15.8%. Given the objective of CBAs and the industries in which they fall, we expect that agreements apply in particular to workers with low levels of skills. Figure 1.1b shows that the share of covered workers is higher at lower percentiles of the wage distribution in 2000 and 2014. Therefore, coverage level drops as skill level rises. In addition, as in most agreements managers are explicitly excluded, coverage for the high-skilled workers is likely to be overestimated by simply looking at industry of employment. Figure A1b offers similar evidence when proxying skill with educational attainment – it is among the tertiary educated that coverage is lowest.

Overall the summary statistics and stylized facts suggest that there is substantial variation across regions in support for immigration and labor market conditions. The objective of this study is to test the extent to which this variation can be attributed to differences in migrant exposure and the level of labor protection across regions.

1.4 Voting Analysis

1.4.1 Votes and Preferences

Our proxy for support for immigration comes from vote outcomes, which is in contrast to the majority of studies that use social survey responses. The benefit of votes is that they show revealed rather than stated preferences. A potential concern with the vote outcomes could be that they do not represent the preferences of the general population. The mean participation rate across the six votes considered is slightly more than fifty percent of eligible voters. Although abstention in single votes can be large, the share of permanent abstainers is estimated to be only between ten and twenty percent in the Swiss context (Sciarini et al., 2016).

An additional concern is that voting results reflect preferences over a specific policy proposal which may not be representative of attitudes towards immigration more generally. This issue is mitigated by including outcomes from several votes. Additionally we rely on the Vox survey data. It offers information on individual-level voting behavior for the same set of votes, and on attitudes. Respondents are asked whether they would prefer Switzerland (1) that gives equal opportunities to foreigners or better chances for the Swiss; (2) that is more open to the outside or more closed. In Table A3 in the Appendix we test whether voting behavior and attitudes are correlated after controlling for individual-level attributes. All

regressions include place of residence and vote fixed effects.[18] Consistently, respondents who state that they are in favor of equal opportunities for foreigners and an open Switzerland are found to be more likely to cast a pro-immigration vote. Therefore, voting behavior is representative of general attitudes towards immigrants.

1.4.2 Immigrant Exposure and Native Voting Behavior

We are motivated by a conceptual framework in which labor market concerns affect support for immigration (see Scheve and Slaughter, 2001). Given that immigrants are overrepresented at the bottom of the skill distribution, we think that labor market concerns are relevant to low-skilled natives. In what follows we test whether such concerns lead to negative voting behavior.

Table 1.2 presents estimates of the impact of a higher foreigner exposure on the share of pro-immigration votes from Ordinary Least Squares (OLS) in Panel A and Instrumental Variable (IV) regressions in Panel B. OLS estimates are insignificant and small in magnitude, while the evidence from the IV specification suggests a positive effect. The magnitude of the latter estimate goes down in the full control specification and becomes insignificant at the ten percent level. A positive effect is in line with the contact hypothesis (Allport et al., 1954).

We modify Equation 1.1 to allow for the direct effect of migrants to depend on native skill levels. We proxy the share of skilled voters using the proportion of upper-secondary or higher educated natives in the municipality based on 2000 census data. In columns (2) and (5) of Table 1.2 skill enters as a continuous variable and in columns (3) and (6) as quintiles to allow for non-linearities. It is evident that as the average skill level in a municipality rises, the response to immigration becomes more positive. To ease interpretation, Figure 1.2a plots estimates from the IV full control specification and shows how a higher foreigner share affects vote outcomes at different quintiles of the skill distribution. Effects are negative for municipalities in the bottom quintile of the distribution, where the share of individuals with at least an upper-secondary degree is less than 61%. An increase in the number of immigrants equal to 1 percent of the native population decreases the share of pro-immigration votes by 0.2 percentage points. At the top of the distribution, where the share of individuals with at least an upper-secondary degree is more than 75%, the effect is opposite – an increase in

[18] Note that place of residence is defined based on a separate classification with sixty-four categories, referred to as agglomerations.

Table 1.2: Voting analysis by native educational level

	Outcome: share of pro-immigration votes					
	Without controls			With controls		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
Sh. migrants	0.018	-0.644	-0.069	-0.020	-0.217	-0.056
	(0.034)	(0.342)	(0.030)	(0.025)	(0.235)	(0.029)
Sh. migrants x sh. skilled		0.933			0.283	
		(0.508)			(0.354)	
Sh. migrants x Q2 sh. skilled			-0.010			0.007
			(0.038)			(0.039)
Sh. migrants x Q3 sh. skilled			-0.011			-0.025
			(0.044)			(0.048)
Sh. migrants x Q4 sh. skilled			0.176			0.117
			(0.061)			(0.057)
Sh. migrants x Q5 sh. skilled			0.145			0.048
			(0.102)			(0.083)
N	13290	13290	13290	13290	13290	13290
Panel B: IV						
Sh. migrants	0.251	-1.341	-0.059	0.109	-1.687	-0.155
	(0.135)	(0.625)	(0.063)	(0.095)	(0.871)	(0.065)
Sh. migrants x sh. skilled		2.068			2.481	
		(0.946)			(1.310)	
Sh. migrants x Q2 sh. skilled			0.013			0.060
			(0.043)			(0.077)
Sh. migrants x Q3 sh. skilled			0.040			0.102
			(0.050)			(0.100)
Sh. migrants x Q4 sh. skilled			0.282			0.318
			(0.075)			(0.128)
Sh. migrants x Q5 sh. skilled			0.316			0.346
			(0.146)			(0.234)
First stage F-stat	23.036	18.728		18.337	9.219	
N	13290	13290	13290	13290	13290	13290

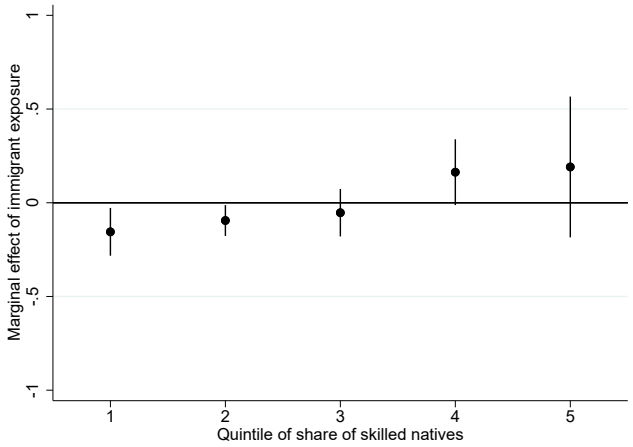
Note: The table presents estimates from OLS and IV regressions using municipality-level data. Share of migrants is the number of foreign residents in a municipality in a year divided by native population in a municipality in 2000. Share skilled is the share of native residents with upper-secondary or higher level of education in 2000. Controls are listed in Table A2; all specifications include municipality and vote fixed effects. Weights assigned to observations equal the number of Swiss residents in 2000. Standard errors in parentheses are clustered at the municipality level. Source: FSO, ZEMIS.

the number of immigrants equal to 1 percent of the native population raises vote shares by close to 0.2 percentage points. Similarly, Figure A3a plots estimates from column (5) against the distribution of the continuous skill measure. While the linearity assumption affects magnitudes, results are qualitatively similar.

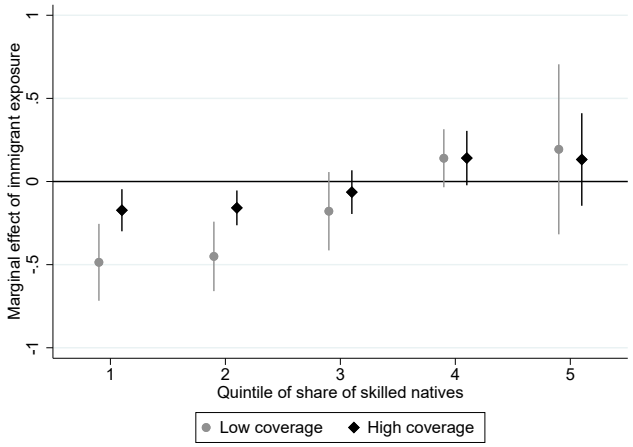
Panel A of Table A4 shows a robustness check where we use the distribution of immigrants across Swiss regions in 1970 when constructing the instrumental variable. The initial distribution of immigrants across the country is unlikely to be correlated with current shocks to either attitudes or labor demand if migrant shares from an earlier year are used. While the first stage becomes weaker and significance levels drop in the second stage, results remain similar to the baseline specification – municipalities with a higher share of low-skilled natives respond

Figure 1.2: Voting analysis by quintiles of native educational attainment

(a) Overall estimates



(b) Estimates by level of CBA coverage



Note: The figure presents estimates from an IV regression using municipality level data. The outcome is the share of pro-immigrant votes. Share of migrants is the number of foreign residents in municipality and year divided by native population in municipality in 2000. Share of skilled natives is the share of native residents with upper-secondary or higher level of education in 2000. Controls are listed in Table A2; all specifications include municipality and vote fixed effects. Weights assigned to observations reflect the number of Swiss residents in 2000. Standard errors are clustered at the municipality level, 95% confidence intervals plotted. In Figure (b) effects at the 10th and 90th percentile of the coverage measure are reported. Source: FSO, SECO, ZEMIS.

more negatively to immigration.

We also find evidence that education level is positively associated with the probability of voting pro-immigration. Pro-immigration vote shares are significantly positively correlated with the skill level in the data from the FSO (correlation coefficient 0.18) and in the survey evidence as presented in Table A3. This pattern and the estimated responses to higher migrant exposure are consistent with a labor market channel where natives who compete against migrants are less in favor of immigration. However, education is likely to affect support for im-

migration through a number of channels. Specifically, it is argued to directly promote tolerance and improve knowledge and appreciation of foreign cultures (see Hainmueller and Hopkins, 2014). Additionally, competition for public goods and services could affect in particular low-skilled natives if they are more likely to use them. If labor protection raises support for immigration for natives who it aims to protect, we consider this as evidence that labor market concerns shape preferences over migration policies. Therefore, our main focus of analysis is on how immigrant presence interacts with collective bargaining coverage in determining vote outcomes.

1.4.3 Collective Bargaining and Vote Effects of Immigration

In the analysis of collective bargaining coverage we follow Equation 1.2. Specifically, Table 1.3 tests the importance of labor protection by introducing a triple interaction between the migrants exposure measure, the proxy for native skill and the share CBA covered. Panel A shows OLS regressions, while Panel B uses an instrumental variable approach for migrant shares and reports reduced form estimates for coverage. In line with our hypothesis, we observe that it is in municipalities with low levels of native educational attainment that CBAs raise pro-immigration vote shares. This holds regardless of whether we use a continuous or a discrete skill measure and is not affected by the inclusion of control variables. Results tend to be stronger in terms of magnitudes in the IV specification relative to the OLS.

In Figure 1.2b we plot the estimates from Panel B column (4). The y-axis shows the marginal effect of a higher foreigner share on vote outcomes at different levels of native skill level and under low versus high levels of coverage. Vote outcomes depend less on the level of skill of the native population under higher levels of CBA coverage. When coverage is low (tenth percentile, i.e. around ten percent coverage), we estimate a significantly negative effect of a higher foreigner presence for municipalities with low native educational attainment. A 1 percent rise in immigration relative to the native population decreases pro-immigration vote shares with 0.5 percentage points. At high levels of coverage (ninetieth percentile, i.e. around twenty percent coverage) the magnitude is much smaller – 0.17 percentage points. At the upper end of the skill distribution, in the forth and fifth quintile, marginal effects are positive and the difference between low and high coverage is insignificant. The difference is alsomarginally insignificant

Table 1.3: Voting analysis by native educational level and CBA coverage

	Outcome: share of pro-immigration votes			
	Without controls		With controls	
	(1)	(2)	(3)	(4)
Panel A: OLS				
Sh. migrants	-1.170	-0.254	-0.701	-0.291
	(0.628)	(0.078)	(0.476)	(0.074)
Sh. migr. x Sh. CBA cov.	2.394	0.687	2.136	0.862
	(1.461)	(0.258)	(1.298)	(0.255)
Sh. migr. x Sh. skilled	1.795		1.080	
	(0.914)		(0.675)	
Sh. migr. x Sh. CBA cov. x Sh. skilled	-3.852		-3.469	
	(2.053)		(1.785)	
Sh. migr. x Q2 sh. skilled		0.064		0.123
		(0.115)		(0.117)
Sh. migr. x Q3 sh. skilled		0.262		0.295
		(0.120)		(0.119)
Sh. migr. x Q4 sh. skilled		0.500		0.502
		(0.099)		(0.087)
Sh. migr. x Q5 sh. skilled		0.451		0.380
		(0.166)		(0.137)
Sh. migr. x Sh. CBA cov. x Q2 sh. skilled		-0.304		-0.458
		(0.332)		(0.350)
Sh. migr. x Sh. CBA cov. x Q3 sh. skilled		-0.984		-1.157
		(0.361)		(0.365)
Sh. migr. x Sh. CBA cov. x Q4 sh. skilled		-1.204		-1.430
		(0.266)		(0.273)
Sh. migr. x Sh. CBA cov. x Q5 sh. skilled		-1.216		-1.304
		(0.359)		(0.336)
N	13290	13290	13290	13290
Panel B: IV				
Sh. migrants	-4.529	-0.646	-4.139	-0.667
	(1.534)	(0.171)	(1.403)	(0.162)
Sh. migr. x Sh. CBA cov.	11.997	1.915	10.124	1.792
	(3.848)	(0.481)	(2.983)	(0.491)
Sh. migr. x Sh. skilled	6.395		5.740	
	(2.217)		(2.044)	
Sh. migr. x Sh. CBA cov. x Sh. skilled	-16.523		-13.707	
	(5.556)		(4.337)	
Sh. migr. x Q2 sh. skilled		0.083		0.048
		(0.163)		(0.185)
Sh. migr. x Q3 sh. skilled		0.405		0.422
		(0.201)		(0.204)
Sh. migr. x Q4 sh. skilled		0.847		0.806
		(0.187)		(0.175)
Sh. migr. x Q5 sh. skilled		1.035		0.896
		(0.308)		(0.344)
Sh. migr. x Sh. CBA cov. x Q2 sh. skilled		-0.317		-0.122
		(0.476)		(0.520)
Sh. migr. x Sh. CBA cov. x Q3 sh. skilled		-1.175		-1.135
		(0.594)		(0.589)
Sh. migr. x Sh. CBA cov. x Q4 sh. skilled		-1.910		-1.786
		(0.544)		(0.543)
Sh. migr. x Sh. CBA cov. x Q5 sh. skilled		-2.651		-2.142
		(0.861)		(0.781)
N	13290	13290	13290	13290

Note: The table presents estimates from OLS and IV regressions using municipality-level data. For specification details see notes to Table 1.2. Source: FSO, SECO, ZEMIS.

for the third quintile. In Figure A3b we present results from column (2) where we use a continuous skill measure. The evidence is consistent.

We conduct a number of robustness checks with respect to the CBA measure. First, we change how agreements are linked to industries. In our baseline specifications we assign each CBA to 3-digit NOGA-08 codes. Alternatively, we can link contracts to more disaggregated industries at the 4-digit level. While this could improve the precision of the treatment assignment, it might give noisier estimates as assigning contracts to narrowly defined industries is ambiguous. This could introduce a measurement error which, if random, results in an attenuation bias.[19] Moreover, large agreements are likely to set standards beyond narrowly defined industries. Results presented in Table A4 are consistent with our baseline findings. Second, we abstract from newly introduced agreements by fixing collective bargaining levels in 2000. This does not alter results qualitatively (results available upon request). Next, we test whether results are robust to the specific votes included. While estimates are generally consistent, including the AuG vote appears crucial for our results (see Table A5).

We conduct our analysis at the municipality level, which relates to the place of living of voters. The benefit of this unit of analysis is the high number of municipalities compared to more aggregated geographical units. On the other hand, voters do not necessarily work in the municipality of living which makes its collective agreements potentially not relevant. Generally valid agreements tend to cover an industry in several municipalities in the same region, so local coverage correlates with coverage in nearby areas. At the same time, low-skilled occupations, which are typically the ones covered by CBAs, are more likely to be locally available than skilled jobs. Therefore, labor protection in the municipality of living is likely to be applicable to the type of workers who are the focus of the study. To alleviate remaining concerns, we run the analysis at the commuting zone level and report the results in Table A6. Significance levels tend to decline compared to the municipality level specification. In line with our baseline results we observe that a higher level of labor market protection increases support for immigration in commuting zones where the native population is relatively low-skilled.

Finally, we look at differences in participation rates between municipalities more and less exposed to immigration. We find some evidence that participation rises as immigration goes up in municipalities at the bottom of the skill distribution and that it falls at the top of the distribution. These effects are the inverse of the vote estimates we previously observed. Furthermore, the response to immigration

[19]The FSO provided the link between collective agreements and industries. Therefore, we do not expect that there is any systematic bias in the measure that could affect estimates in a specific direction.

tends to be more positive in magnitude in municipalities with higher coverage levels. As this effect is consistent across skill levels of the native population, it cannot explain our results.

We have provided evidence that labor protection is linked to a more positive response to immigration and argue that this is in line with a conceptual framework in which individual labor market concerns shape voting behavior. In the next section we test how native labor market outcomes respond to immigration and whether this response depends on the extent to which native workers are covered by collective agreements.

1.5 Labor Market Analysis

1.5.1 Immigrant Exposure and Native Wage Outcomes

In the main analysis of wage outcomes we follow Dustmann et al. (2013) and proxy returns to skill with percentiles of the native wage distribution. Table 1.4 presents estimates at the 50th, 5th, 10th and 95th percentile while Figure 1.3a shows full control IV results at every 5th percentile of the native wage distribution. It is evident that all estimates below the 30th percentile are negative and significant at the five percent level. Individuals at the tenth percentile are the most negatively affected where an increase in immigration equal to 1 percent of the native population leads to a 0.7 percent decrease in wages. At the 5th percentile the estimate is slightly lower in absolute magnitude indicating a drop of 0.6 percent. Estimates among the very high-skilled individuals, on the other hand, are positive though too noisy to reject the null hypothesis of no effect. Including controls alters the magnitude of the estimates while results remain qualitatively unchanged. The more negative IV estimates compared to those from the OLS regressions suggest that migrants, particularly those with low to medium skill levels, sort themselves into areas with improving economic conditions. Overall, our results are in line with predictions from a classic labor market model where an inflow of workers in a skill group reduce wages in the same group.

Dustmann et al. (2013) show that under a nested constant elasticity of substitution (CES) production function, estimated parameters at each percentile of the wage distribution are proportional to the density of immigrants at that point. Assuming that migrant location in the distribution is constant across regions and years, and capital is perfectly elastic in supply, one can combine the point estimates with the density of migrants in the wage distribution and derive the elasticity of substitution across skill types. Figures 1.1a and A2 show that dif-

Table 1.4: Wage analysis by native percentiles of the wage distribution

	Outcome: ln **real gross hourly wage at the m-th percentile**			
	50th pct	5th pct	10th pct	95th pct
	(1)	(2)	(3)	(4)
Panel A: OLS				
Sh. migrants	-0.110	-0.240	-0.273	0.138
	(0.093)	(0.105)	(0.101)	(0.174)
Mean outcome	3.553	3.062	3.157	4.224
Sd outcome	0.103	0.070	0.068	0.195
N	848	848	848	848
Panel B: OLS with controls				
Sh. migrants	-0.077	-0.218	-0.242	0.165
	(0.070)	(0.095)	(0.086)	(0.159)
Mean outcome	3.553	3.062	3.157	4.224
Sd outcome	0.103	0.070	0.068	0.195
N	848	848	848	848
Panel C: IV				
Sh. migrants	-0.056	-0.625	-0.701	0.493
	(0.092)	(0.241)	(0.241)	(0.339)
Mean outcome	3.553	3.062	3.157	4.224
Sd outcome	0.103	0.070	0.068	0.195
First stage F-stat	9.056	9.056	9.056	9.056
N	848	848	848	848
Panel D: IV with controls				
Sh. migrants	-0.082	-0.628	-0.697	0.472
	(0.153)	(0.250)	(0.261)	(0.347)
Mean outcome	3.553	3.062	3.157	4.224
Sd outcome	0.103	0.070	0.068	0.195
First stage F-stat	9.722	9.722	9.722	9.722
N	848	848	848	848

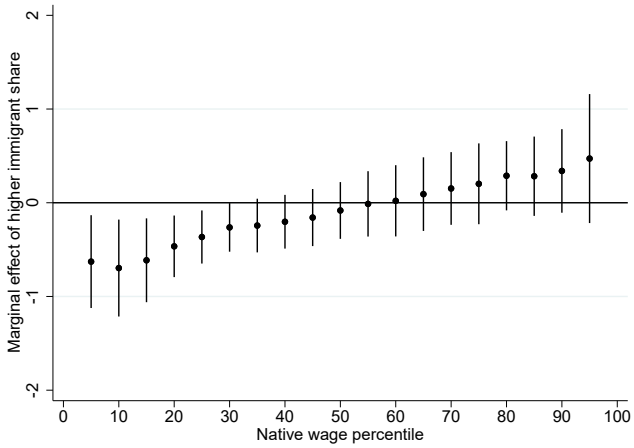
Note: The table presents estimates from OLS and IV regressions using biennial data at the commuting zone level between 2000 and 2014. Share of migrants is the number of foreign residents in a commuting zone in a year divided by native population in a commuting zone in 2000. Controls are listed in Table A2; all specifications include commuting zone and year fixed effects. Weights assigned to observations equal the number of natives employed in commuting zone in 2000. Standard errors in parentheses are clustered at the commuting zone level. Source: FSO, SESS, ZEMIS.

ferences across time and geographical areas exist, but trends are overall fairly similar. Therefore, we find no evidence that the necessary assumptions behind the methodology are not satisfied in our context. Indeed, results shown in Figure 1.3a vary inversely with the position of immigrants in the wage distribution presented in Figure 1.1a. We compute the position of migrants along the native wage distribution for the whole period and calculate an elasticity of substitution across labor types of approximately 1.1. Compared to a value of 0.6 for the UK estimated in Dustmann et al. (2013), wages in the Swiss context appear to be less responsive to a labor supply shock in the same skill group.[20]

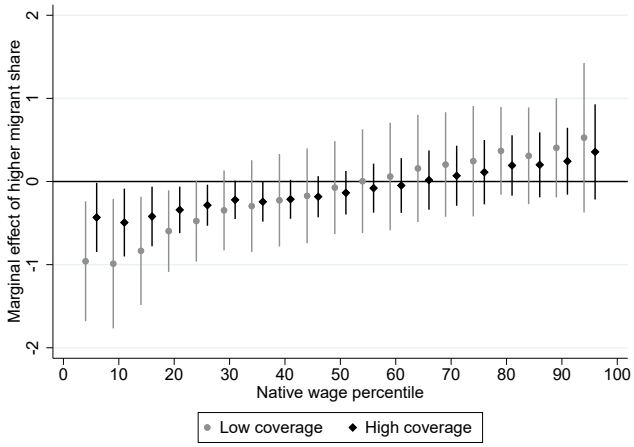
[20]Note that the rank sensitivity assumption is satisfied under the estimated elasticity of substitution (see p.149-150 of Dustmann et al., 2013)

Figure 1.3: Wage analysis by native percentiles of the wage distribution

(a) Overall estimates



(b) Estimates by level of CBA coverage



Note: The figure presents estimates from IV regressions using biennial data at the region level between 2000 and 2014. The outcome is the ln real gross hourly wage at the m-th percentile. Share of migrants is the number of foreign residents in region and year divided by native population in region in 2000. Effects at the 10th and 90th percentile of the coverage measure are reported. Controls are listed in Table A2; all specifications include commuting zone and year fixed effects. Weights assigned to observations equal the number of natives employed in commuting zone in 2000. Standard errors are clustered at the commuting zone level, 95% confidence intervals plotted. In Figure (b) effects at the 10th and 90th percentile of the coverage measure are reported. Source: FSO, SECO, SESS, ZEMIS.

Next we report a number of robustness checks. In Table A7 Panel A we construct the instrumental variable using 1970 instead of 2000 census data on the distribution of immigrants across Swiss regions. Table A8 Panel A presents similar evidence when we separate workers based on educational attainment. While average effects are insignificant, they mask considerable differences across skill groups. Individuals with lower- and upper-secondary education tend to be negatively affected while effects in the group of tertiary educated workers are positive and not significant. Results are, therefore, consistent with the main estimates.

1.5.2 Collective Bargaining and Wage Effects of Immigration

We have documented that low-skilled natives are more likely to be employed in industries with collective bargaining agreements. Low-skilled natives are also negatively affected by immigration in the labor market. Therefore, we test whether the magnitude of the wage effects depends on the level of labor protection. Table 1.5 shows our main results where skill is proxied with percentiles of the native wage distribution. Panels A and B report OLS estimates, while Panels C and D report IV results. In both types of specifications, we find evidence that a higher coverage mitigates the negative effects of immigration in particular for the very low-skilled workers – those at the fifth and tenth percentiles of the wage distribution. The interaction term is also significant at the fifteenth percentile and at the twentieth percentile it becomes marginally insignificance at the ten percent level. Adding controls does not affect the size of the estimates significantly.

To give some intuition about the magnitude of these effects, in Figure 1.3b we show how wages at every fifth percentile respond to immigration under low and high levels of coverage. Under higher coverage levels the effects of migration are generally less pronounced. At high levels of coverage at the bottom of the skill distribution marginal effects go down in absolute terms and are close to 0.4–0.5 percent. At low levels of coverage, on the other hand, marginal effects are close to -1.

We have so far provided evidence that CBAs introduce a rigidity in wages. How is the elasticity of substitution across skill groups affected by different coverage levels? Following the same methodology as before and using parameter estimates reported in Figure 1.3b, we obtain an elasticity of 0.8 at low and 1.65 at high levels of CBA coverage. As the level of wage rigidity goes up with collective bargaining coverage, wages are less responsive to a relative labor supply shock. This by construction translates into a higher elasticity. Therefore, the estimate of a production function parameter is dependent on the level of labor market protection. This evidence is consistent with the work of Foged et al. (2019) who argue that estimates of wage elasticities obtained from immigration studies conducted in different countries reflect the nature of labor market institutions at destination.

In the Appendix we confirm that wage results are robust to changes in how we construct the CBA coverage measure (see Table A7). Results remain unchanged – collective agreements unambiguously improve native wage outcomes under higher

Table 1.5: Wage analysis by native percentiles of the wage distribution and CBA coverage

	Outcome: ln **real gross hourly wage at the m-th percentile**			
	50th pct	5th pct	10th pct	95th pct
	(1)	(2)	(3)	(4)
Panel A: OLS interaction				
Sh. migrants	-0.212	-0.447	-0.454	0.128
	(0.120)	(0.120)	(0.119)	(0.261)
Sh. migr. x Sh. CBA cov.	0.362	0.939	0.872	-0.119
	(0.198)	(0.329)	(0.333)	(0.509)
Mean outcome	3.553	3.062	3.157	4.224
Sd outcome	0.103	0.070	0.068	0.195
N	848	848	848	848
Panel B: OLS interaction with controls				
Sh. migrants	-0.088	-0.375	-0.349	0.230
	(0.094)	(0.108)	(0.094)	(0.250)
Sh. migr. x Sh. CBA cov.	0.001	0.726	0.560	-0.403
	(0.223)	(0.308)	(0.264)	(0.507)
Mean outcome	3.553	3.062	3.157	4.224
Sd outcome	0.103	0.070	0.068	0.195
N	848	848	848	848
Panel C: IV interaction				
Sh. migrants	-0.029	-1.415	-1.452	0.647
	(0.270)	(0.474)	(0.496)	(0.608)
Sh. migr. x Sh. CBA cov.	-0.358	4.605	4.516	-1.241
	(1.323)	(1.712)	(1.725)	(2.074)
Mean outcome	3.553	3.062	3.157	4.224
Sd outcome	0.103	0.070	0.068	0.195
First stage F-stat	3.483	3.483	3.483	3.483
N	848	848	848	848
Panel D: IV interaction with controls				
Sh. migrants	-0.015	-1.454	-1.454	0.689
	(0.462)	(0.558)	(0.604)	(0.678)
Sh. migr. x Sh. CBA cov.	-0.555	4.711	4.426	-1.535
	(1.778)	(1.999)	(2.090)	(2.402)
Mean outcome	3.553	3.062	3.157	4.224
Sd outcome	0.103	0.070	0.068	0.195
First stage F-stat	3.638	3.638	3.638	3.638
N	848	848	848	848

Note: The table presents estimates from OLS and IV regressions using biennial data at the commuting zone level between 2000 and 2014. Share of migrants is the number of foreign residents in a commuting zone in a year divided by native population in a commuting zone in 2000. Controls are listed in Table A2; all specifications include commuting zone and year fixed effects. Weights assigned to observations equal the number of natives employed in commuting zone in 2000. Standard errors in parentheses are clustered at the commuting zone level. Source: FSO, SECO, SESS, ZEMIS.

immigration rates. Similarly, using a time-invariant coverage measure does not affect results qualitatively. Finally, we run the analysis using educational attainment as a proxy for skill in Table A8. Wage percentiles are particularly attractive in the analysis of collective agreements which are likely relevant to those at the bottom of the skill distribution. We show that the mitigating effect of CBAs holds for native workers with up to lower-secondary education who are likely at

the bottom of the wage distribution. Results are, thus, robust to the skill proxy used.

1.5.3 Immigrant Exposure and Native Employment Outcomes

Wage effects are of first-order interest given the nature of the labor protection we study. Another possible margin of adjustment is employment. For instance, firms may dismiss employees as a response to increasing labor costs caused by the collective bargaining clauses. In Table 1.6 we investigate employment effects. Our most stringent specifications in which we instrument for migrant presence and include the full set of controls (see Panel D) show an insignificantly negative overall effect. Upper-secondary educated natives, however, have lower employment rates in regions with a higher migrant exposure. On the other hand, we see no significant change in employment among the lower-secondary educated, i.e. those individuals who experience a drop in wages. We interpret results as driven by differences in the labor supply elasticity across skill groups (Dustmann et al., 2017). Table 1.7 shows that collective agreements mitigate the negative employment effects for the group of upper-secondary educated natives. Robustness checks with respect to an alternative instrument and CBA measure are presented in Table A9 and confirm the baseline results. Overall the evidence we provide for labor market effects is consistent with the findings of Edo and Rapoport (2019) who study the response of native labor market outcomes under different levels of minimum wages in the United States.

Table 1.6: Employment analysis by native educational attainment

	Outcome: share of natives employed in population 18-65			
	All	Up to lower-secondary	Upper-secondary	Tertiary
	(1)	(2)	(3)	(4)
Panel A: OLS				
Sh. migrants	-0.260	-0.361	-0.224	-0.073
	(0.068)	(0.250)	(0.079)	(0.084)
Mean outcome	0.776	0.451	0.787	0.909
Sd outcome	0.047	0.117	0.059	0.053
N	1590	1576	1590	1585
Panel B: OLS with controls				
Sh. migrants	-0.166	-0.215	-0.226	-0.092
	(0.061)	(0.256)	(0.078)	(0.083)
Mean outcome	0.776	0.451	0.787	0.909
Sd outcome	0.047	0.117	0.059	0.053
N	1590	1576	1590	1585
Panel C: IV				
Sh. migrants	-0.683	-0.385	-0.583	-0.313
	(0.306)	(0.444)	(0.270)	(0.233)
Mean outcome	0.776	0.451	0.787	0.909
Sd outcome	0.047	0.117	0.059	0.053
First stage F-stat	7.730	7.710	7.730	7.721
N	1590	1576	1590	1585
Panel D: IV with controls				
Sh. migrants	-0.418	0.050	-0.609	-0.390
	(0.259)	(0.420)	(0.277)	(0.252)
Mean outcome	0.776	0.451	0.787	0.909
Sd outcome	0.047	0.117	0.059	0.053
First stage F-stat	7.451	7.424	7.451	7.444
N	1590	1576	1590	1585

Note: The table presents estimates from OLS and IV regressions using annual data at the commuting zone level between 2000 and 2014. Share of migrants is the number of foreign residents in a commuting zone in a year divided by native population in a commuting zone in 2000. Controls are listed in Table A2; all specifications include commuting zone and year fixed effects. Lower-secondary level of education is compulsory education as highest degree, upper-secondary is an apprenticeship or a matura, tertiary is a degree from a university, university of applied sciences, university of teacher education or a professional degree. Weights assigned to observations equal the number of native respondents 18-65 years of age in commuting zone in 2000. Standard errors in parentheses are clustered at the commuting zone level. Source: FSO, SLFS, ZEMIS.

Table 1.7: Employment analysis by native educational attainment and CBA coverage

	Outcome: share of natives employed in population 18-65			
	All	Up to lower-secondary	Upper-secondary	Tertiary
	(1)	(2)	(3)	(4)
Panel A: OLS interaction				
Sh. migrants	-0.248	-0.522	-0.170	-0.017
	(0.111)	(0.519)	(0.131)	(0.172)
Sh. migr. x Sh. CBA cov.	0.024	0.718	-0.127	-0.192
	(0.434)	(2.160)	(0.495)	(0.679)
Mean outcome	0.776	0.451	0.787	0.909
Sd outcome	0.047	0.117	0.059	0.053
N	1590	1576	1590	1585
Panel B: OLS interaction with controls				
Sh. migrants	-0.204	-0.476	-0.180	-0.022
	(0.126)	(0.522)	(0.133)	(0.170)
Sh. migr. x Sh. CBA cov.	0.239	1.134	-0.109	-0.246
	(0.512)	(2.268)	(0.490)	(0.662)
Mean outcome	0.776	0.451	0.787	0.909
Sd outcome	0.047	0.117	0.059	0.053
N	1590	1576	1590	1585
Panel C: IV interaction				
Sh. migrants	-1.509	-1.319	-1.143	-0.673
	(0.721)	(1.067)	(0.573)	(0.602)
Sh. migr. x Sh. CBA cov.	4.538	3.836	3.378	1.790
	(2.462)	(4.194)	(1.937)	(2.232)
Mean outcome	0.776	0.451	0.787	0.909
Sd outcome	0.047	0.117	0.059	0.053
First stage F-stat	3.225	3.211	3.225	3.222
N	1590	1576	1590	1585
Panel D: IV interaction with controls				
Sh. migrants	-1.058	-0.424	-1.198	-0.831
	(0.617)	(0.978)	(0.579)	(0.650)
Sh. migr. x Sh. CBA cov.	3.393	1.155	3.556	2.269
	(2.105)	(3.923)	(1.935)	(2.363)
Mean outcome	0.776	0.451	0.787	0.909
Sd outcome	0.047	0.117	0.059	0.053
First stage F-stat	3.142	3.128	3.142	3.140
N	1590	1576	1590	1585

Note: The table presents estimates from IV regressions using annual data at the commuting zone level between 2000 and 2014. Share of migrants is the number of foreign residents in a commuting zone in a year divided by native population in a commuting zone in 2000. Controls are listed in Table A2; all specifications include commuting zone and year fixed effects. Lower-secondary level of education is compulsory education as highest degree, upper-secondary is an apprenticeship or a matura, tertiary is a degree from a university, university of applied sciences, university of teacher education or a professional degree. Weights assigned to observations equal the number of native respondents 18-65 years of age in commuting zone in 2000. Standard errors in parentheses are clustered at the commuting zone level. Source: FSO, SECO, SLFS, ZEMIS.

1.6 Conclusion

We examine how exposure to migrants affects native support for immigration and labor market outcomes. Our results show that subgroups of natives that are

negatively affected in terms of labor market outcomes are less in favor of loose immigration regulation. Support for immigration, however, rises under stronger labor protection for these subgroups. Our results are consistent with labor market concerns shaping preferences over alternative immigration policies. At a time of rising anti-migrant sentiments, this study contributes to a debate on what determines attitudes towards foreigners. Importantly, our findings have implications for the type of policies which alleviate economic concerns and increase support for immigration.

The need for social protection in the broader context of globalization has been emphasized in the literature (Rodrik, 1997). We add to the discussion of policies which affect attitudes towards immigration by assessing the role of labor market protection. This paper shows that CBAs can be effective in boosting support for immigration by preserving labor market standards. Compliance with the negotiated clauses can be imperfect, however, and reduce their effectiveness. Switzerland and other countries such as Austria have introduced supporting measures to strengthen the enforcement of labor protection at the same time as removing restrictions to immigration. The general policy lesson of our findings is that setting common labor market standards within industries and effectively enforcing them raises support for immigration by mitigating labor market concerns.

Our labor market results measure short-term effects. Capital adjustments, incentives to switch occupations and acquire more skills likely offset any short-term effects of migrant inflows. Similarly, labor regulation could slow down such adjustments and, thus, affect long-term native wage and employment outcomes. For example, D'Amuri and Peri (2014) provide evidence that natives are less likely to switch their occupations following an immigrant inflow if employment regulation is stricter. If such adjustments occur in the longer run, the evidence offered in this study is not indicative for how CBAs affect labor market outcomes after markets have adjusted. Any policy recommendations are, thus, conditional on the weight one puts on alleviating current labor market concerns and raising support for immigration in the short-run.

2 Free Movement of Workers and Native Demand for Tertiary Education

with Mirjam Bächli

2.1 Introduction

Governments play an important role in shaping access to education and often aim at achieving specific national educational targets. However, other factors are also relevant in the decision to accumulate human capital. For example, labor market reforms can alter relative wages and are therefore likely to have an impact on demand for schooling. Ultimately, such changes affect the composition of human capital and, thus, the economy's potential to innovate and grow (Lucas, 1988). In this paper, we focus on demand for degrees acquired at the tertiary level of education. We investigate whether changes in labor market conditions induced by an inflow of foreign workers affect enrolment and study field choice of natives. This question is especially relevant for today's knowledge-based societies in a context of free movement of workers.

We study a major Swiss migration reform that abolished quotas and introduced free movement of workers between Switzerland and the member states of the European Union (EU) and the European Free Trade Association (EFTA). The Agreement on the Free Movement of Persons (AFMP) was signed in 1999 and approved by the electorate in 2000. The proposed changes were in particular important for cross-border commuters, i.e. individuals who work in Switzerland but reside abroad. Restrictions on commuting were gradually relaxed during the transition period 2002 to 2007 until their complete abolishment. As a result, the number of frontier workers substantially increased. Their share relative to total employed rose from 4% in 2001 to 5.8% in 2015. These values understate the commuters' relevance for border regions, where their share was 13.7% in 2015.

The Swiss context offers advantages in addition to the exogenous policy change. Since cross-border commuters reside abroad, they leave demand for goods and services in the country of work largely unaffected. Importantly, they are unlikely to demand publicly provided services such as education. The inflow of commuters can, therefore, be regarded as an almost pure labor supply shock. The country's dual education system gives access to tertiary education to graduates from general training at Universities and from vocational training at Universities of Applied

Sciences. These two groups of students are interesting to compare as they differ in their opportunity costs of studying and in their outside options. Importantly, the Swiss education system enables us to isolate education demand from supply forces since fulfilling the admission requirements generally guarantees enrolment.

In our identification strategy, we distinguish between affected and non-affected labor market regions based on driving distance to the national border. We combine this cross-sectional variation in exposure with the timing of the reform implementation in a difference-in-differences framework. Our main outcome of interest is native educational enrolment at the tertiary level by institutional type and study field. We look at regional labor market outcomes by educational level to examine incentives to accumulate human capital. Enrolment data come from administrative sources and labor market outcomes from large-scale surveys.

The analysis of the policy change reveals an increase in the share of cross-border commuters of 6 percentage points in the treated relative to the control regions in the post-reform period. This is driven by upper-secondary and tertiary educated commuters. Therefore, labor market competition increases in particular for natives who are at the point of deciding whether to join the labor market or enrol in tertiary education. We observe that commuters are overrepresented in STEM occupations and expect this to play a role in the choice of study field.

Our results show that natives in regions affected from cross-border commuting respond by demanding more tertiary education relative to natives from regions less affected. Enrolment in undergraduate degrees from universities of applied sciences rises in the post-reform period in treated regions by 1.8 percentage points. This effect is statistically significant and economically large relative to the mean enrolment rate of 13%. In line with this finding, we estimate a drop in wages of upper-secondary educated and a rise in wages of tertiary educated workers. These changes mean lower opportunity costs to studying for upper-secondary graduates and higher returns to tertiary degrees. Upper-secondary graduates for whom the labor market is relevant are mainly those with a vocational training. A majority of them joins the labor market immediately after completing the apprenticeship but this becomes less attractive as an option due to the fall in upper-secondary wages. As a possible response vocationally trained can gain access to tertiary education through universities of applied sciences, which is indeed what we see in the data. We do not observe a difference in the evolution of enrolment at universities between treatment and control regions. We explain this result by the fact that the very purpose of a general education is enrolment in tertiary education.

Furthermore, we map occupations to fields of study using survey data and measure the extent to which specific fields are affected by the labor supply shock. Subjects are considered to be affected if they are linked to occupations that commuters hold relatively more often than resident workers. Those coincide with STEM (science, technology, engineering, and mathematics) fields. We find that enrolment in less affected fields of study rises in the post-reform period in treated regions. These are non-STEM subjects that often require skills tied to the country-specific context compared to STEM fields.

Our results add to different strands of the literature. First, we contribute to the research on educational choice that aims to understand what motivates individuals to acquire tertiary education (see Altonji et al., 2016, for a summary). There is a general consensus that expected earnings matter (e.g., Beffy et al., 2012; Wiswall and Zafar, 2015) but also some evidence of limited knowledge of returns (Xia, 2016). We find that enrolment in tertiary education responds to changes in returns to schooling but field choice does not. A closely related literature links enrolment decisions to changes in opportunity costs of studying as induced by business cycle fluctuations. There is evidence that enrolment is countercyclical in lower educational levels (Ayllon and Nollenberger, 2016), in college (Dellas and Sakellaris, 2003; Long, 2014) and in graduate school for women (Johnson, 2013). The majority of these studies use unemployment rates to measure business cycle fluctuations. We test how changes in wages and unemployment induced by an inflow of frontier workers affect undergraduate enrolment decisions at the tertiary level. We distinguish between students with a vocational and general background who are likely to respond differently to changing labor market returns to education. While recessions could have additional implications on budget cuts and therefore on the supply of education, immigration reforms are unlikely to have such indirect effects.

A number of studies link native demand for education to immigration. An inflow of foreign students can affect school resources while foreign workers may change labor market returns to education. Betts (1998) finds an overall negative effect of immigrant inflows on high school graduation rates of American-born minorities. Hunt (2017) finds that a higher share of low-skilled adult immigrants has a positive impact on high-school completion, while immigrants of school age have no significant effect. Similarly for college enrolment, Jackson (2018) shows a significant positive impact of labor immigrants but no effect of foreign students in the cohort. Focusing on labor immigrants, McHenry (2015) documents an increase in native post-secondary degree attainment. Llull (2018) considers education, participation and occupation as margins of adjustments to immigration. Educa-

tional attainment depends on relative changes in wages, which in turn depend on the capital adjustments in the economy. A subset of this literature investigates how an increase in the number of international students affects native educational decisions at the tertiary level. Shih (2017) finds an increase in domestic enrolment at US universities, Machin and Murphy (2017) in postgraduate studies in the UK, while Borjas (2004) shows crowding-out effects for white men at US elite universities. Our contribution is to differentiate native tertiary enrolment according to institutional types. In addition, we isolate the effect of a labor supply shock by focusing on cross-border commuters, who do not compete with natives for school resources.

Studies document that foreign-born workers are more often employed in scientific and technical occupations than natives (Hunt and Gauthier-Loiselle, 2010; Peri and Sparber, 2011; Hanson and Slaughter, 2016). Related research links migrant employment into occupations to native enrolment in specific study fields. Ransom and Winters (2020) find an outflow of native-born Americans, specifically blacks, from STEM subjects related to occupations with more foreign workers. Cortés and Pan (2015) document a similar crowding-out effect from nursing studies. We add to this literature by considering native enrolment in all study fields and by grouping fields by the intensity of the labor supply shock.

Related research looks at demand for specific education fields as a response to immigrant students. Orrenius and Zavodny (2015) find that demand for STEM fields is lower among women in age cohorts with a higher share of foreign-born students. Similarly, Anelli et al. (2018) show that a higher share of foreign-born students in the introductory math course lowers the probability that natives will graduate with a STEM degree, while the subset of international students has a positive effect. While we are interested in how foreign workers affect native educational enrolment, we test whether our study field results can be explained by crowding-out effects of foreign students. We do not find evidence in favor of this hypothesis.

We further build on the migration literature which finds mixed evidence on the impact of immigrant labor on native wages (see e.g., Borjas, 2003; Ottaviano and Peri, 2012b; Dustmann et al., 2016b). A subset of the studies focuses on cross-border commuting. Looking at the same reform, Beerli et al. (2018) show a positive effect of relaxing restrictions on the inflow of foreign workers on the wages of high-skilled natives and no negative employment effects. Dustmann et al. (2017) investigate a temporary increase in low-skilled Czech cross-border commuters into Germany after the fall of the Berlin wall. They find a negative

impact on native wages and in particular on employment levels. We complement this literature by studying how policy induced changes to labor market returns affect incentives for human capital formation.

The remainder of the article is organized as follows. In Section 2.2 we discuss the Swiss context with the regulatory framework applied to cross-border commuters and the educational system. In Section 2.3 we describe the data and outline the empirical strategy. In Sections 2.4 and 2.5 we present our results on enrolment in tertiary education and on enrolment by field of study. In Section 2.6 we conclude.

2.2 Context

2.2.1 Policy Change on Cross-Border Commuting

Cross-border commuters with a citizenship from a European Union (EU) or European Free Trade Association (EFTA) member state working in Switzerland are subject to the rules outlined in the Agreement on the Free Movement of Persons (AFMP). It was signed in June 1999, approved by the electorate in May 2000 and introduced on the 1st of June 2002. While the agreement affects all foreign workers from EU and EFTA countries, we focus on cross-border commuters. Cross-border commuters are non-Swiss by nationality and require a G-permit to work in Switzerland. Since they need a working contract from a Swiss employer to receive or extend such a permit, the number of cross-border commuters consists of foreign workers only.

Prior to the AFMP, cross-border commuters and firms that wanted to hire them had to fulfil several requirements. Commuters had to live in formal border zones in the neighboring countries. Within Switzerland, they were only allowed to work in defined border zones. Permits were tied to a specific employer and valid for up to one year after which they had to be renewed. Commuters had to return to their place of origin on a daily basis. Furthermore, employers had to prove that the vacancy could not be filled by a native worker (local priority requirement).

The policy change was implemented in three steps. After June 2002 cross-border commuters were free to reside outside the border zones of the home country. In addition, they were required to return to their place of residence only once a week rather than every day. The work permit was no longer bound to a specific job and its validity was extended to the length of the working contract, for a maximum of five years. In June 2004 the local priority requirement was abolished and as a result, cross-border commuters could be hired under the same conditions as

resident workers in the Swiss border zones. Full liberalization across the entire country came into force in June 2007 when all restrictions on cross-border working were removed. In particular, cross-border commuters from countries within the EU-17 and EFTA were allowed to work anywhere in Switzerland, while interim regulations applied for other EU member states.

The new rules on the free movement of cross-border commuters led to a large increase in the number of foreign workers. The majority of them commutes into border regions because commuting costs increase with distance. In these regions, the share of commuters in total employed rose from 9.9% in 2001 to 13.7% in 2015. In 2017, 95% of all cross-border commuters are nationals of France, Germany, Italy or Austria. They generally commute to municipalities in which their mother tongue is spoken, suggesting that language skills matter. 97–98% of the Austrian and German commuters work in a municipality in which German is spoken by the majority of residents. The share of Italian and French commuters that go to Italian- and French-speaking municipalities is 88% and 80% respectively.

Earnings survey data show that in 2016 48% of cross-border commuters have an upper-secondary degree, 23% up to a lower-secondary degree, 19% an academic tertiary and 10% a professional tertiary degree. In comparison, the share of native workers with a lower secondary education is significantly lower (11%) and the share with upper-secondary significantly higher (59%). These differences persist during the study period. The share of tertiary educated workers has been rising for both native and frontier workers since 1996 but the trend is stronger for the latter group. Given that the focus of the study is academic tertiary education, we abstract from professional tertiary degrees.

Data from the Federal Statistical Office show that in the period 1999–2017 a considerable share of cross-border commuters are employed in manufacturing compared to the industry's share in total employment (26% compared to 14%). In particular, the production of computer and electronic, chemical and pharmaceutical goods is highly reliant on frontier workers. Since 2006, commuters are overrepresented in administrative and support services relative to industry size. Throughout the period commuters are underrepresented in industries with a high level of government involvement such as education (3% compared to 7% in total employment), health care and social work (10% compared to 14%), public administration and defence (1% compared to 4%).

2.2.2 Dual Education System

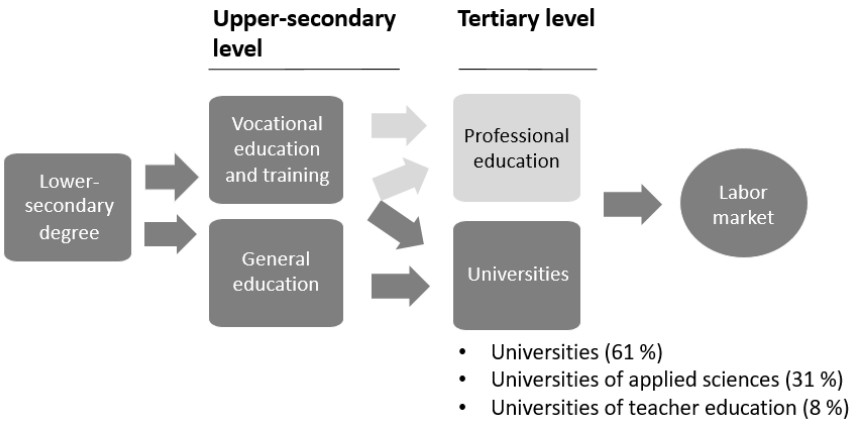
Three types of tertiary education institutions exist in Switzerland: Universities and Federal Institutes of Technology, Universities of Applied Sciences, Universities of Teacher Education. Universities and the Federal Institutes of Technology (UNI) are the oldest institutions with a right to grant tertiary level degrees. In 1995 the foundation for a dual tertiary education system was laid with the introduction of the vocational matura at the upper-secondary level. This allowed the establishment of Universities of Applied Sciences (UAS) in 1997. While universities are committed to a combination of teaching and research, universities of applied sciences impart professional skills with a practice and application oriented focus. The difference is apparent in how the two types of institutions select professors. Professors at a UAS require several years of professional experience and expert knowledge in the field. Universities, on the other hand, look for professors with a purely academic career. Teacher education has belonged to the tertiary level since 2001. Some parts of the country have independently organized Universities of Teacher Education (UTE), whereas in others teacher education is taught at universities of applied sciences or at universities. Of all tertiary students in the academic year 2017/2018, 61% are enrolled in a university, followed by 31% in universities of applied sciences and the remaining 8% in universities of teacher education.

Both UNI and UAS offer STEM and non-STEM education. Around 69% of all university students in the year 2017/2018 are enrolled in a non-STEM field. At universities of applied sciences this share is almost 74%. In the first years after their foundation, the UAS specialized in STEM education. Over time they heavily expanded to non-STEM teaching. The UNI and the UAS teach courses that are offered by both institutions and also exclusively by only one. As a consequence, the introduction of the UAS increased the available study fields and hence the range of skills that can be acquired through higher education.

Different educational pathways give access to tertiary education. Figure 2.1 shows that at the upper-secondary level one can follow a vocational or a general education track. According to the Swiss Federal Statistical Office 68.3% of students in upper-secondary education pursued a vocational degree in 2016, while the rest were enrolled in general education. Depending on the chosen track of vocational education, an apprentice can graduate with a diploma, tailored for joining the labor market, or gain in addition a vocational matura. Such a matura can be obtained during or after the vocational training. It is required for admission to a university of applied sciences. A general education results in either a general

or a specialised matura. The general matura grants access to universities and universities of teacher education, but can also be used to enter a university of applied sciences. The specialised matura has both general and vocational education components and grants access to universities of applied sciences and universities of teacher education. In 2016, 21.2% of the Swiss residents under the age of 25 hold a general, 15.4% a vocational, and 3% a specialised matura.

Figure 2.1: Swiss education system



Note: Three different types of universities with different admission requirements exist at the tertiary level of education. Of all enrolled students in the academic year 2017/2018, 61% are at a university, 31% at a university of applied sciences and the remaining 8% at a university of teacher education. Source: SHIS-studex.

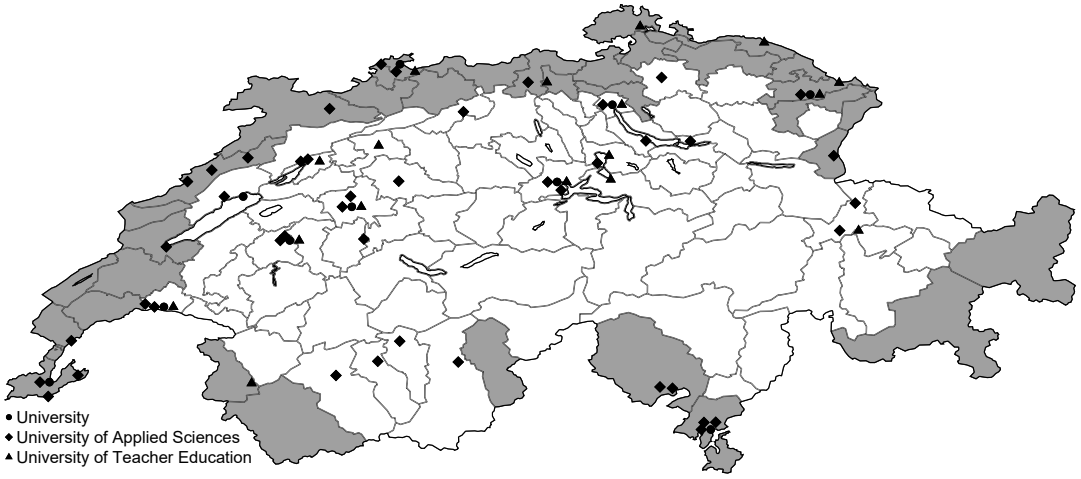
Figure 2.2 shows the locations of the tertiary education institutions across Switzerland in 2017. Most of the institutions are in the northern and western part of the country and clustered in the main centres. There are ten cantonal universities and two federal institutes of technology spread over ten cities. In contrast, most of the nine universities of applied sciences have several locations, which are often specific to a study field. Finally, there are twenty institutions that offer teacher education. The high density of institutions enables daily commuting to classes for a large share of the population. This allows living with parents while studying and, therefore, reduces study costs.[1] Semester fees for Swiss nationals are generally below CHF 1,000 and make up a small part the study costs.

Besides a degree requirement, no major supply side entry restrictions exist for Swiss residents at the undergraduate level. A general matura typically grants access to any degree in the chosen university.[2] As an exception, health degrees can have a cap on the number of students enrolled in a year. To enrol in a specific degree, universities of applied sciences can require a certain thematic focus of the

[1]The yearly costs of living are estimated by study advisory services to be around CHF 23,000 (e.g., University of Zurich).

[2]This does not hold for foreign students without a Swiss matura (12.5% of all first-year students in the academic year 2017/2018) for which explicit quotas may be imposed.

Figure 2.2: Locations of tertiary institutions



Note: This map shows Switzerland's 106 commuting zones split into treated (grey) and control regions (white). The locations of the tertiary institutions in 2017 are shown by institutional type.

vocational matura or relevant work experience. Interviews are often conducted to test the ability of candidates in social or health related fields at UAS. While there is little screening when entering a tertiary institution, the pool of eligible students is already selective by the admission requirements for upper-secondary education tracks resulting in a matura. Furthermore, graduation rates are below unity with 85% of those who enrolled in a bachelor program in 2007 and having completed their education by 2017. Therefore, in the analysis we will look at both enrolment and graduation rates.

2.3 Data and Methods

2.3.1 Data

We combine several data sources to conduct our analysis.

In the analysis of the educational choice we use administrative data referred to as SHIS-studex, an abbreviation for the Swiss Higher Education Information System. This is an individual-level database covering all matriculated students at the academic tertiary level of education in Switzerland. The data is provided by the Federal Statistical Office (FSO). It includes students at universities since 1990, universities of applied sciences since 1997, and universities of teacher education since 2001. The variables used are age, nationality, place of residence prior to beginning a study, certificate granting access to tertiary education, type of tertiary institution and field of study. The structure of the SHIS-studex dataset allows tracking individuals from the point of enrolment up to graduation and

provides information on received degrees.

We are interested in demand for undergraduate degrees and focus on first-year students enrolled in a bachelor study over the period 1997–2017. We select students who completed their matura in Switzerland in order to assign them to the region of residence at the time of receiving the certificate. Additionally, we focus on Swiss nationals because they are likely to be more familiar with the choice set in a dual tertiary education system compared to non-Swiss. To calculate the share of students enrolled we divide the number of first-year students by the birth cohort size. The cohort is the Swiss population in each region at the median age of first-year students in year 1997. In the full sample the median age is 21, in the sample of students enrolled in universities it is 20 and in universities of applied sciences and universities of teacher education 22. FSO provides information about the size of the native population at the municipality level and age structure of the population at the canton level.

We add to the SHIS-studex dataset information from the Survey of Higher Education Graduates (EHA) which has been conducted biennially since 2003. It has a panel structure where individuals respond to questions related to their working experience and acquired skills one and five years after graduation. Additionally, it is linked to administrative information from the SHIS-studex dataset about the participants' tertiary education. In the first-wave survey, all graduates who have successfully completed a degree in a Swiss institution of tertiary education in the previous year are asked to complete a questionnaire. Only graduates who participated in the first-wave survey are asked to take part in a second-wave survey four years later. Our focus lies on first-wave results because we are interested in information collected a short time after graduation. We consider the subset of Swiss graduates with a bachelor or master who have in addition a Swiss matura. We use information about place of living, place of work and the mapping between fields of study and occupations.

The Swiss Earnings Structure Survey (SESS) is a large-scale firm survey conducted biennially in the month of October since 1994. It is a repeated cross-section of private sector firms in the secondary and tertiary sectors of the economy. We use information on the region in which the firm is located and its industry. Companies provide information on a random subset of employees. The number of workers covered depends on the firm size, with data for at least one third of all workers. The sample is limited to employees 18–65 years of age, with available gross hourly wages, region of work, permit type, gender and education. Working permit information distinguishes native from cross-border employees.

We differentiate three types of education based on the highest level attained – tertiary, upper-secondary and up to lower-secondary training. Our analysis covers the years from 1996 to 2016. We use data on native gross hourly wages and on the share of cross-border commuters. For the latter measure we divide the number of commuters by the number of total employees in 1996. In the analysis by educational level, the share of cross-border commuters is the number of commuters by education divided by the total number of employees in 1996, independent of the educational level.

While the SESS covers only employed individuals, the Swiss Labor Force Survey (SLFS) includes individuals aged 15 years and older. The survey has been conducted annually in the second quarter of the year from 1996 to 2009 and quarterly from 2010. For consistency, we use annual data. In 2018 almost 65,000 individuals were surveyed. Information about municipality of residence, demographic characteristics, educational attainment and employment outcomes for the household head is available. We limit the sample to individuals in the age group 18–65. Unemployment is defined as not being employed, but searching and being available for a job. Employment is defined as being employed for a salary, by a family member or self-employed. Students, retired individuals and people inactive for other reasons are considered to be out of the labor force. The native unemployment rate is the number of unemployed relative to total labor force by educational category. The native employment rate is the number of employed relative to total number of individuals by educational category.

Additionally, we obtain the travel time data for each municipality from www.map.search.ch, which we accessed in December 2018. We measure the distance by car from each municipality to the closest border crossing or border checkpoint according to the Federal Customs Office. The travel time for a region is the time weighed by total number of employed in 1995 in each municipality within a region. Regions with a border crossing or border checkpoint are assigned a value of zero minutes.

2.3.2 Empirical Strategy

Motivated by the nature of the policy change, the empirical analysis is based on a standard difference-in-differences strategy. We investigate the reform effects by comparing regions close to the border with those further away before and after

the regulatory change. To focus on local labor markets, we take the MS-region[3] or commuting zone – short region – as unit of observation. Regions are constructed based on commuting patterns and are, therefore, considered small-scale labor markets. The allocation of municipalities to the total 106 regions rests on 2000 census data and is provided by the FSO. We follow Beerli et al. (2018) and use a fixed threshold to define the treated group. Out of the 106 regions, 35 are less than or equal to thirty minutes of travel time away from the national border and are assigned to the treatment group as shown in Figure 2.2. We consider the remaining 71 regions to be too far from the border to attract cross-border workers and hence assign them to the control group. While this is consistent with the commuting pattern observed in the data, we also use alternative thresholds in sensitivity checks.

We run the following specification

$$\begin{aligned} y_{rt} = \alpha + \beta_1 \text{Transition}_t \times 1(Dist_r \leq 30min) + \\ \beta_2 \text{Post}_t \times 1(Dist_r \leq 30min) + \mathbf{X}'_{rt}\gamma + \delta_r + \varepsilon_{rt} \end{aligned} \tag{2.1}$$

where r is region, and t year. First-year students are allocated to their region of residence at the time of taking the matura, while in the labor market analysis individuals are assigned to the region of the workplace. The observation period 1997–2017 is split into pre-reform (1997–2001), transition (2002–2007) and post-reform (2008–2017) periods. The outcome y_{rt} measures either the enrolment rate or labor market conditions. The coefficients of interest, β_1 and β_2, show the difference in the dependent variables between treated and control regions during and after the reform compared to pre-reform years.

In our baseline specification we include region fixed effects to capture time-invariant regional variation in the outcomes of interest and we limit the control variables to NUTS II region $\times$ year fixed effects. The latter control for changes over time occurring at the larger geographical level.[4] In the enrolment analysis, we also control for the natural log of native population that may drive enrolment rates. Further variables that could vary during the period and across regions are introduced in robustness checks of the analysis of educational outcomes. We use weights to account for the different employment and population sizes across re-

[3]MS comes from the French "mobilité spatiale". According to commuting data from the FSO, 63% of employees work and live in the same commuting zone in 2014. The equivalent unit in the USA is also called commuting zone.

[4]Switzerland has seven NUTS II regions, each containing between one and seven cantons. Cantons are the largest administrative sub-national units, followed by districts and municipalities. The education system is organized on a cantonal level, while a tertiary institution's catchment area often extents over several cantons.

gions. In the regressions on enrolment we weigh by native cohort size in 1997, in the wage analysis by the number of native employees in 1996, in the regressions on unemployment rates by the labor force in 1996, and in the case of employment rates by the total number of individuals in 1996. Finally, the labor supply shock regressions are weighed by total employment in 1996. In a robustness check we confirm that the weights do not drive our results. Standard errors are clustered at the regional level.

While β_1 and β_2 are the only estimates we report in tables, graphically we present the results from an event study. The coefficients capture the effect of the reform in year t relative to the last year in the pre-reform period.

$$y_{rt} = \alpha + \sum_{t=1997}^{2017} \beta_t Year_t \times 1(Dist_r \leq 30min) + \mathbf{X}'_{rt}\gamma + \delta_r + \varepsilon_{rt} \tag{2.2}$$

The key assumption under which our results are valid is that enrolment rates and labor market conditions would have followed the same trend in treatment and control regions absent the reform. We compare yearly coefficients as visualized in the figures to investigate whether this assumption is likely to hold. We find no evidence that the common trends assumption is violated. Similarly, results are robust to adding additional control variables which could have evolved differently over time in the treatment and control regions. These results are reported in more detail in Section 2.4.1.

The Stable Unit Treatment Value Assumption (SUTVA) is the second important precondition to be fulfilled. We argue that spillovers between the treatment and control group do not play a significant role. Such spillovers could emerge if, for instance, individuals from the treatment and control group attend the same high-school and are taught about labor market competition and outcomes in their common surrounding. First, our unit of analysis is the commuting zone which is the relevant local labor market given the place of residence prior to enrolment. Second, information frictions naturally make someone more aware of local employment conditions, especially at a young age. Third, we know from the EHA survey where former students work and live and can compare these locations with the one where they grew up. In 2017 59% of the graduate students live one year after graduating in the same region where they resided during their upper-secondary education. 29% even work in that same region – a considerable share given that many high-skill jobs are not available throughout the country. Fourth, our sample consists of natives with a Swiss entry exam. Natives are likely

to perceive the local labor market conditions as more important than foreign residents, who may also consider opportunities abroad or be internally more mobile (Schündeln, 2014).

2.3.3 Summary Statistics

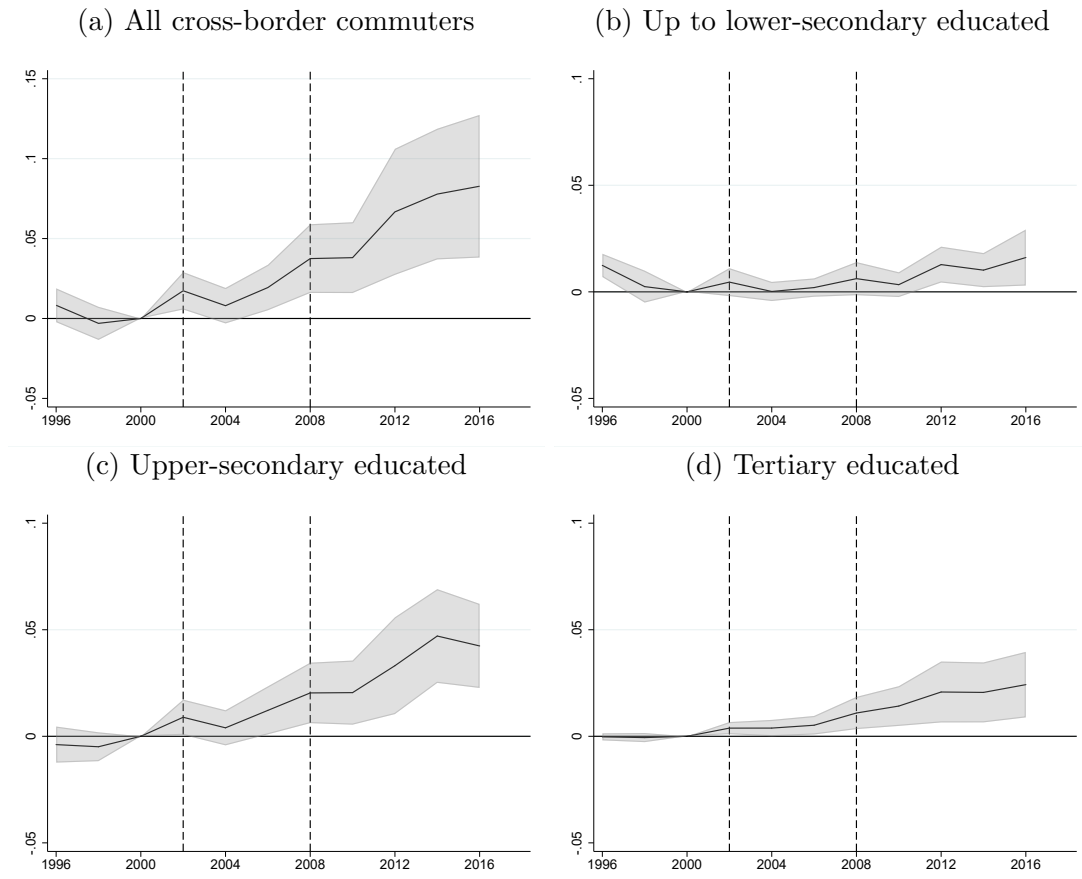
In this section we present summary statistics distinguishing treatment and control regions. To justify the treatment assignment rule, we estimate Equation 2.1 and compare the share of cross-border commuters across treatment and control regions in the different periods. Column (1) of Table 2.1 shows that regions within thirty minutes of travel time from the national border experienced a positive labor supply shock relative to regions further away, in particular in the post-reform period. The magnitude of the effect is almost 6 percentage points – an impact that is both statistically significant and economically relevant. Figure 2.3a presents the size and timing of the inflow of commuters for each year. Magnitudes increase after the second implementation step of the AFMP in 2004 from 0.8 to 8.3 percentage points in 2016. Figure B1a replicates these results with administrative data. In line with survey data, we find increasing effects from the transition period onwards. Administrative data shows that cross-border commuting was already slightly on the rise in the last years of the pre-treatment period. This could be explained by the informal relaxation of migration regulations prior to 2002. We take this into account when discussing the timing of the enrolment results.

Table 2.1: Exposure to cross-border commuters by educational level

	Outcome: share of cross-border commuters			
	All	Up to lower-secondary	Upper-secondary	Tertiary
	(1)	(2)	(3)	(4)
30min * 2002-2007	0.013	-0.003	0.011	0.005
	(0.006)	(0.002)	(0.004)	(0.002)
30min * after 2008	0.059	0.005	0.036	0.018
	(0.017)	(0.004)	(0.009)	(0.006)
Mean outcome	0.070	0.020	0.038	0.012
Sd outcome	0.115	0.053	0.056	0.021
Commuting zones	106	106	106	106
within 30 min	35	35	35	35
N	1166	1166	1166	1166

Note: This table shows difference-in-differences estimates using biennial data at the commuting zone level for the period 1996–2016. Outcome is the share of cross-border commuters in total employment. Denominator is fixed at first year. Observations are weighed by number of total employees in first year. Standard errors in parentheses are clustered at the commuting zone level. Sources: SESS.

Figure 2.3: Exposure to cross-border commuters



Note: This figure shows difference-in-differences estimates using biennial data at the commuting zone level for the period 1996–2016. The vertical lines indicate the beginning of the transition period (2002) and of the post-treatment period (2008). Outcome is the number of cross-border commuters divided by total employment. Denominator is fixed at first year. Observations are weighed by number of total employees in 1996. Standard errors are clustered at the commuting zone level, 95% confidence intervals shown. Source: SESS.

In Table 2.1, columns (2)–(4), and Figures 2.3c and 2.3d we find that the labor supply shock consists of upper-secondary and tertiary educated commuters. More than half of the total increase of 6 percentage points reflects a rise in upper-secondary educated commuters. The share of tertiary educated commuters grows by 1.8 percentage points in the post-reform period. Results show no significant increase in commuting of lower-secondary educated workers during the period.

In the Appendix we present robustness checks. In Table B1 we test the sensitivity of the results to lower and higher cut-off values.[5] We find that the estimated magnitude of the supply shock declines as we choose a higher threshold value. As a generalization, we use a continuous measure for travel time. The treatment is defined as an exponential function of travel time, taking into account the non-linear link between share of commuters and a region's distance from the border. The continuous measure confirms that the inflow of commuters is increasing over time in regions which are closer to the border. These results assure us that the chosen baseline threshold, which is motivated by observed commuter flows, is robust to alternative definitions of the treatment variable. A last concern we address is whether resident migrants are like the commuters more often employed in border regions. Figure B1b shows that the share of resident migrants does not evolve differently across treatment and control regions during the study period. We therefore focus on cross-border commuters as the relevant group of foreign workers.[6]

The summary statistics in Table 2.2 reveals further differences and similarities across treatment and control groups. The average share of first-year students is 41% in the treated and 36% in the control regions. This is driven by enrolment at universities while rates are similar for the other two institutional types. Although enrolment rates vary across subjects and regions, the relative attractiveness of the study fields is similar among the two groups with the exception of teacher training. Average wages are comparable in both groups and there is a considerable wage premium for tertiary graduates relative to upper-secondary educated workers. The share of unemployed natives in the treatment group is 3.4% compared to 2.7% in the control group. Similarly, employment is on average higher in

[5] According to data from the FSO on commuting patterns of the resident population, around 68% of the commuting workers need thirty minutes or less to reach the place of work in 2010. Although cross-border commuters are excluded from this data, we do not have any indication that their average travel time should differ largely.

[6] According to individual level migration data (ZEMIS) provided by the State Secretariat of Migration (SEM), the share of cross-border commuters that switched from a G-permit to a resident permit between 2002 and 2018 is 15.5%. The robustness test performed alleviates concerns about a potentially determining role of former cross-border commuters in the distribution of resident migrants across treatment and control regions.

the control region. These differences hold for each of the three education groups.

Table 2.2: Summary statistics

	Treatment group			Control group		
	N	Mean	Sd	N	Mean	Sd
Share of cross-border commuters	385	0.162	0.138	781	0.010	0.016
... with lower-secondary education	385	0.048	0.076	781	0.002	0.007
... with upper-secondary education	385	0.088	0.062	781	0.005	0.008
... with tertiary education	385	0.026	0.028	781	0.002	0.004
Share enrolled	735	0.411	0.127	1491	0.355	0.109
... at UNI	735	0.237	0.103	1491	0.191	0.071
... at UAS	735	0.145	0.056	1491	0.134	0.050
... at UTE	625	0.032	0.023	1236	0.035	0.019
... in agriculture	735	0.004	0.003	1491	0.004	0.003
... in arts and humanities	735	0.044	0.020	1491	0.035	0.015
... in business and law	735	0.108	0.039	1491	0.093	0.033
... in education	735	0.043	0.021	1491	0.042	0.022
... in engineering	735	0.057	0.019	1491	0.054	0.018
... in health	735	0.054	0.036	1491	0.040	0.027
... in ICT	735	0.013	0.007	1491	0.013	0.007
... in math and sciences	735	0.037	0.014	1491	0.033	0.013
... in services	735	0.005	0.006	1491	0.004	0.005
... in social sciences	735	0.043	0.023	1491	0.035	0.017
Mean ln gross hourly wage	385	3.574	0.102	781	3.563	0.109
... of lower-secondary educated	385	3.295	0.087	781	3.298	0.086
... of upper-secondary educated	385	3.522	0.083	781	3.498	0.081
... of tertiary educated	385	3.935	0.086	774	3.936	0.086
Share unemployed	735	0.034	0.022	1491	0.027	0.018
... with lower-secondary education	730	0.070	0.082	1354	0.055	0.077
... with upper-secondary education	735	0.035	0.026	1491	0.028	0.023
... with tertiary education	692	0.025	0.027	1445	0.017	0.023
Share employed	735	0.758	0.051	1491	0.786	0.046
... with lower-secondary education	735	0.445	0.117	1433	0.467	0.129
... with upper-secondary education	735	0.768	0.063	1491	0.799	0.057
... with tertiary education	711	0.889	0.057	1446	0.917	0.051

Note: Share of cross-border commuters is in 1996 total employment. Lower-secondary level of education is compulsory education as highest degree, upper-secondary is an apprenticeship or a matura, tertiary is a degree from a university, university of applied sciences or teacher education. Share enrolled is number of first-year students divided by cohort size in 1997. UNI is short for university, UAS for university of applied sciences and UTE for university of teacher education. One-digit ISCED fields of studies are considered. Share unemployed is number of unemployed divided by labor force. Share employed is the number of employed divided by the population of working age. Weights assigned to the observations reflect number of native employees in 1996, native cohort size in 1997, number of total employees in 1996, native labor force in 1996, and native population of working age in 1996. All data is at the commuting zone level. The observation period for the enrolment outcomes is 1997–2017 and for the other outcome variables 1996–2016, respectively. Sources: SESS, SLFS, SHIS-studex.

2.4 Enrolment in Tertiary Education

2.4.1 Demand for Tertiary Degrees

We begin with examining whether regions more affected by the introduction of the free movement of cross-borders commuters have different enrolment rates relative to regions less affected. Results in Table 2.3 show a rise in overall enrolment

in the post-reform period among individuals residing in the affected region prior to beginning their studies. The magnitude of the effect is 2.6 percentage points, which is equivalent to 7% relative to the mean enrolment rate. The analysis by institutional type in columns (2)–(4) indicates that individuals from regions close to the border experienced a rise in enrolment at universities of applied sciences. The magnitude of the effect is 1.8 percentage points or approximately 13% relative to the mean of the outcome.[7] In contrast, we find no evidence that entry into universities and universities of teacher education differs between the treatment and the control regions in any of the periods.[8] For the pre-reform period, Figure 2.4 shows that there is no significant difference in enrolment between the treatment and control group. This evidence suggests that the parallel trend assumption is not violated. Indeed, the timing of the increase in enrolment is in line with the intensity of the labor supply shock presented in Figure 2.3a. While we observe a small increase in commuting prior to 2002, we find that enrolment goes up only in the post-reform period when all barriers were abolished and the inflow of frontier workers was substantial.

Table 2.3: Enrolment by institutional type

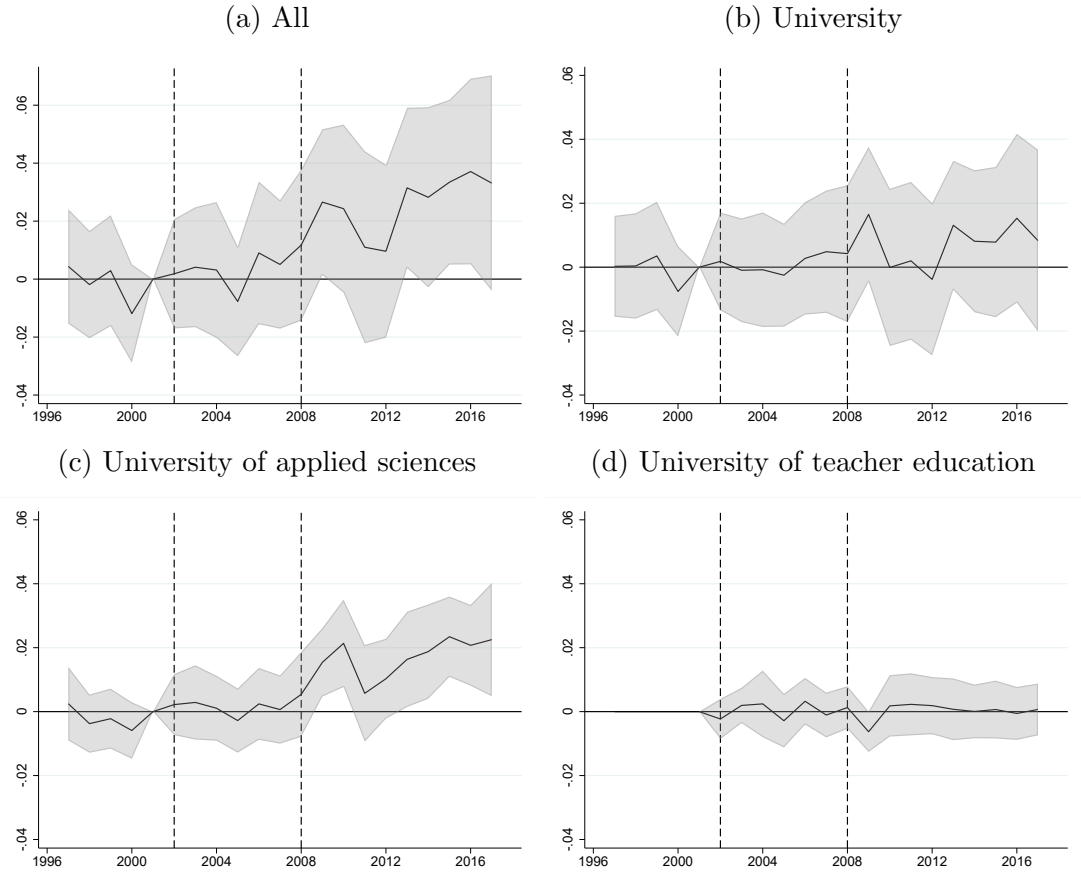
	Outcome: share of enrolled native first-year students			
	All	University	University of applied sciences	University of teacher education
	(1)	(2)	(3)	(4)
30min * 2002-2007	0.004	0.002	0.003	0.000
	(0.007)	(0.005)	(0.004)	(0.003)
30min * after 2008	0.026	0.008	0.018	0.000
	(0.011)	(0.008)	(0.005)	(0.004)
Mean outcome	0.372	0.207	0.136	0.035
Sd outcome	0.119	0.086	0.052	0.020
Commuting zones	106	106	106	106
within 30 min	35	35	35	35
N	2226	2226	2226	1802

Note: This table shows difference-in-differences estimates using annual data at the commuting zone level for the period 1997–2017. Outcome is the share of native first-year students in birth cohort. Denominator is fixed at first year and specific to the education category. Observations are weighed by cohort size in a specific education category in first year. Standard errors in parentheses are clustered at the commuting zone level. Source: SHIS-studex.

[7]The range of the coefficient is between 1.4 and 2.1 percentage points when leaving out a labor market region one by one. All results are statistically significant at the 5% level. Note that there are sixteen labor market regions covering between two and eighteen commuting zones.

[8]For university enrolment as an exception, data is available from the early 1990s. In a setting with an extended pre-reform window from 1992–2001, we find no statistical differences between treatment and control regions over all years. Consistently, when leaving out a labor market region one by one the range of the coefficient for university enrolment after 2008 is between 0.2 and 1.4 percentage points and the coefficients are in fifteen out of sixteen cases not statistically significant at the conventional levels. In one exception the coefficient becomes marginally significant at the 10% level.

Figure 2.4: Enrolment by institutional type



Note: This figure shows difference-in-differences estimates using annual data at the commuting zone level for the period 1997 – 2017. The vertical lines indicate the beginning of the transition period (2002) and of the post-treatment period (2008). Outcome is the share of native first-year students in birth cohort. Denominator is fixed at first year and specific to the education category. Observations are weighed by cohort size in a specific education category in first year. Standard errors are clustered at the commuting zone level, 95% confidence intervals shown. Source: SHIS-studex.

In the Appendix we provide a number of robustness checks. Panel A of Table B2 shows that the threshold of thirty travel minutes is not decisive for the main results. The estimates remain similar when using the continuous measure for travel time (Panel B) or if we run the specification at the municipality level (Panel C). The latter shows that there may be an accompanying rise in enrolment in universities but this is of lower magnitude than the university of applied sciences result and gives noisy estimates in an event study.

Table B3 investigates whether our results are sensitive to additional control variables and the weighting scheme. Changes in the supply of education and demand for labor could be confounding factors to the common trend assumption if they differ over time and across regions. Since our observation period coincides with the expansion of the UAS, we test whether enrolment rates are driven by the

availability of new study locations and study fields.[9] Column (2) shows that results are robust to controlling for the presence of tertiary institutions as well as the number of study fields offered within a radius of 20km from the largest municipality in a region in 1990. Labor demand could confound the results if regions closer to the international market face a different trend in their labor demand than regions in the inner part of Switzerland. We proxy labor demand with a Bartik type measure of employment, relying on the industrial composition of each region in 1995 and aggregate annual employment growth at the industry level (see Bartik, 1991, for an initial application to labor demand).[10] As shown in column (3), controlling for labor demand does not change results compared to our baseline specification. Additionally, in column (4) we confirm that weights do not drive the results.

2.4.2 Mechanism

Returns to education To explain the relative rise in demand for tertiary education in treated regions, we look at changes occurring in the labor market. If the policy change has persistent effects on local labor market outcomes, incentives to study change. We expect that lower opportunity costs of studying or better employment prospects for those with a tertiary degree drive enrolment in tertiary education. We consider wages, unemployment and employment rates as labor market outcomes of interest.

We consider the inflow of commuters as an almost pure labor supply shock because their consumption is concentrated in the country of residence and not country of work. Following a classic labor market model with fixed labor demand, we expect a fall in the wages of native workers in increasing competition with foreign workers. Table 2.4 shows a decrease in wages for upper-secondary educated workers and an increase in wages for tertiary educated workers in affected regions, with statistically significant effects in the post-treatment period. In Figure B2 we focus on the evolution of wages across regions and periods. The wage pressure on upper-secondary educated started in the early transition years. In contrast, tertiary educated see an upward trend in their wages in the post-treatment period, while these results are imprecisely estimated. This evidence is

[9] Hoxby (2009) finds for the USA that university choice is less driven by distance partly due to lower transportation costs. In the context of Switzerland, Denzler and Wolter (2011) argue that the distance to university matters for both the decision to enrol and the study field choice in particular for individuals from middle and low socio-economic groups.

[10] We construct the variable as follows: $Bartik_{rt} = \sum_i Sh\,employed_{ir1995} \times \frac{Nr\,Employed_{it}}{Nr\,Employed_{i1995}}$, where i denotes industry, r region and t year. The industry is measured by two-digit NOGA-08 codes.

in line with the results reported in Beerli et al. (2018) who investigate the Swiss labor market consequences of the introduction of the free movement reform. The authors explain the rise in wages for tertiary degrees by an increase in the labor demand of skill-intensive incumbent and new firms. This can lead to higher innovation, productivity or capital formation in a setting with increasing returns to high-skilled labor.[11]

Table 2.4: Wages by educational level

	Outcome: ln **gross hourly wage rate of natives**			
	All	Up to lower-secondary	Upper-secondary	Tertiary
	(1)	(2)	(3)	(4)
30min * 2002-2007	-0.007	-0.018	-0.011	0.018
	(0.008)	(0.012)	(0.008)	(0.011)
30min * after 2008	-0.010	-0.011	-0.012	0.035
	(0.007)	(0.016)	(0.006)	(0.016)
Mean outcome	3.567	3.297	3.504	3.936
Sd outcome	0.106	0.083	0.082	0.086
Commuting zones	106	106	106	106
within 30 min	35	35	35	35
N	1166	1166	1166	1159

Note: This table shows difference-in-differences estimates using biennial data at the commuting zone level for the period 1996–2016. Outcome is the mean natural log of gross hourly wage of natives in a region. Observations are weighed by number of native employees in a specific education category in first year. Standard errors in parentheses are clustered at the commuting zone level. Source: SESS.

Switzerland is split into four language regions.[12] The inflow of commuters in the post-reform period is significant in the German and French speaking regions and larger in the latter. Since the number of units with the main spoken language Italian or Romansh is small, we do not investigate them separately. We perform a leave-one-out analysis by language region to test whether wage effects and enrolment responses are similar (results available upon request). Overall, the reported results are not driven by a single region. Dropping German speaking regions, however, tends to reduce significance levels as it decreases the sample to 31 regions only. The negative effect on wages for upper-secondary educated workers is robust in magnitude to leaving out any region. Excluding the French

[11] Our framework deviates from Beerli et al. (2018) in at least two respects that may explain the different magnitude of the wage effect on tertiary educated natives. First, we use 2000 as the reference year in our event study analysis, while they take 1998. Second, in our measure for tertiary educated we only include individuals with an academic degree, while they also consider individuals with professional tertiary degrees. Our analysis leads to the same qualitative results as theirs.

[12] In 75 out of 106 regions the majority speaks German, in 23 French and in 8 either Italian or Romansh. Within the treated regions, the French speaking regions (eleven) and the Italian speaking regions (three) are overrepresented while the German speaking regions are underrepresented (twenty). There are only two regions with the main language Romansh, whereas one is treated.

speaking regions reduces the magnitude of the wage increase for tertiary educated, while excluding German speaking Switzerland raises it. These results suggest that the wage effect on high-skilled is driven by the French speaking Switzerland. Dropping one language region at a time has little effect on enrolment results. The increase at universities of applied sciences is apparent in all specifications. University enrolment rises when leaving out the German speaking regions but this effect is imprecisely estimated.

In Table 2.5 we look at native unemployment rates. Our analysis shows that rates do not statistically differ over time between treated and control regions for any of the three educational levels. In spite of the large inflow of cross-border commuters, natives in affected regions are not pushed into unemployment. This result suggests that there is some skill complementarity between native and foreign employees. We also look at native employment rates across affected and non-affected regions. Results in Table B4 do not show statistically significant effects.

Table 2.5: Unemployment rates by educational level

	Outcome: native unemployment rate			
	All	Up to lower-secondary	Upper-secondary	Tertiary
	(1)	(2)	(3)	(4)
30min * 2002-2007	0.001	-0.015	0.003	0.000
	(0.003)	(0.016)	(0.004)	(0.005)
30min * after 2008	0.003	-0.005	0.003	0.004
	(0.003)	(0.014)	(0.004)	(0.005)
Mean outcome	0.030	0.062	0.031	0.020
Sd outcome	0.017	0.064	0.020	0.020
Commuting zones	106	103	106	103
within 30 min	35	35	35	34
N	1166	1122	1166	1132

Note: This table shows difference-in-differences estimates using biennial data at the commuting zone level for the period 1996–2016. Outcome is the number of native unemployed divided by the total labor force in an education category. Observations are weighed by the labor force in a specific education category in the first year. Standard errors in parentheses are clustered at the commuting zone level. Source: SLFS.

Given the lack of evidence on (un)employment, we consider changes in wages and therefore returns to education as the main channel driving higher enrolment. Another complementary mechanism is the observed rise in labor market competition. When measuring the level of competition by the inflow of commuters in the same education category, individuals with upper-secondary training are the ones most affected by the reform. Whether it is the decrease in wages or the increase in labor market competition that drives the enrolment results depends on the visibility of the two factors.

Labor market experience Why does enrolment only increases in universities of applied sciences and not in universities? Students at tertiary institutions have different educational backgrounds that expose them to different opportunities on the labor market. The majority of students at a university have a general education while at a university of applied sciences students typically have a vocational training. Numbers from FSO for 2012 graduates show that 64% of those with a vocational matura enrol in tertiary education within 42 months after graduation. This is significantly lower compared to 94% with a general and 84% with a specialised matura (Strubi et al., 2018). The labor market is thus relevant as an outside option for the vocationally trained, while the objective of a general training is to prepare for enrolment at university. Consistently, there are around 3% natives with a matura on the labor market in 2016, while the share of those with an apprenticeship is around 58%. Furthermore, vocationally trained individuals have at least three years of work experience and are therefore likely to be informed about labor market conditions in their field.[13] Based on these numbers, we expect individuals with a vocational matura to respond stronger to changes in labor market prospects than individuals with a general training.

We test our hypothesis by splitting the first-year students at a UAS by their certificate granting access to tertiary education. A vocational matura can be completed during the vocational training (Type I), or in two to four semesters after the vocational education (Type II). A smaller number has either a general or a specialised matura. The three kinds of matura have distinct curricula, resulting in different labor market experiences and opportunity costs of studying. Table 2.6 illustrates that the higher demand for tertiary education is driven by people who do their vocational matura at the same time as their vocational education or have a specialised matura. Columns (1) and (2) reveal differences in the impact of the inflow of foreign workers on first-year students with a vocational matura completed during and after the apprenticeship, respectively. To better understand the differential response we look at the fields in which the two groups of upper-secondary graduates gained their labor market experience. A vocational matura is offered in six major fields. These fields are typically closely related to the vocational education and hence to the occupation in which the apprentice is trained. From the pooled SHIS-studex dataset over the period 1997–2017 we know that more than 90% of the UAS first-year students have a vocational

[13] In contrast, students are likely to be less familiar with change in returns to tertiary degrees. We hence expect that individual perceptions about future earnings deviate from the earnings data used in our analysis (Jensen, 2010; Schweri and Hartog, 2017). This makes the observed decline in upper-secondary wages a more accurate driving force of tertiary enrolment compared to tertiary wage changes.

matura in the three major fields: economics, business and services; technology, architecture and life sciences; health and social work. A larger share of first-year students with a Type I matura has a background in economics, business and services (48.9%) compared to those with a Type II matura (36.7%). The latter group, in contrast, is relatively more often educated in occupations related to health and social work (13.2% versus 3.6%). This descriptive evidence suggests that the higher sensitivity to labor supply shocks of Type I matura students could be linked to their stronger background in market-driven occupations such as business administration or retail.

Table 2.6: Enrolment at UAS by type of entry exam

	Outcome: share of enrolled native first-year students			
	Vocational matura (during)	Vocational matura (after)	Specialised matura	General matura
	(1)	(2)	(3)	(4)
30min * 2002-2007	0.000	0.001	0.002	-0.001
	(0.002)	(0.001)	(0.002)	(0.001)
30min * after 2008	0.008	0.000	0.008	-0.002
	(0.004)	(0.003)	(0.002)	(0.002)
Mean outcome	0.049	0.037	0.012	0.023
Sd outcome	0.026	0.023	0.015	0.014
Commuting zones	106	106	106	106
within 30 min	35	35	35	35
N	2226	2226	2226	2226

Note: This table shows difference-in-differences estimates using annual data at the commuting zone level for the period 1997–2017. Outcome is the share of native first-year students at universities of applied sciences in birth cohort. Denominator is fixed at first year. Observations are weighed by cohort size in first year. Column (1) shows first-year students with a vocational matura done during the apprenticeship, column (2) first-year students with a vocational matura done after the apprenticeship. Standard errors in parentheses are clustered at the commuting zone level. Source: SHIS-studex.

Selection Individuals who are induced to enrol in tertiary education by the reform can be positively or negatively selected (see Cortés and Pan, 2015, for a positive selection into the nursing study). We investigate whether academic achievement as an indicator for student quality differs between the treatment and control regions. We compute the graduation rates of students and allocate them to the first year of enrolment. In Table B5 we find no significant differences between treated and control regions. The higher demand for tertiary degrees in affected regions is driven by students with an average quality similar to that in control regions. Hence, the increase in enrolment leads to a higher share of tertiary graduates coming from affected regions. As a degree is considered to be a key signal for high ability, our evidence suggests that those who responded to the reform on average improve their labor market prospects (Arrow, 1973).

2.5 Enrolment by Field of Study

2.5.1 Affected Study Fields

The enrolment analysis has shown that natives respond to the inflow of frontier workers by demanding more tertiary education. A further margin of adjustment for those who enrol is the choice of study field. In this section we investigate the extent to which specific study fields are affected by the free movement reform. We start by linking subjects to occupations and create the variable *Sh empl*$_j$ which reflects the share of employees trained in a field j.

$$\text{Sh empl}_j = \sum_{o=1}^{O} \text{Sh empl in o} \times \text{Sh empl in o with j}, j \in [1, 22] \tag{2.3}$$

Sh empl in o with j is the share of employed in an occupation o with a degree in field j, which we multiply with the share of employed in the same occupation *Sh empl in o*. Intuitively, we allocate individuals employed in an occupation to fields of study and take into account the size of the occupation.

We infer the link between study fields and occupations from their joint distribution provided by the EHA survey (2003–2017). This approach is consistent with the fact that natives do not observe the education of commuters but have some knowledge of their occupations. We use the study fields at the two-digit ISCED level with twenty-two categories and consider ten high-skilled occupations, defined as ISCO-08 level 1 (managerial occupations) and level 2 (professional occupations). FSO administrative data provides the distribution of cross-border commuters across occupations in 1999 – the first year with available occupation information – while census data from 2000 offers information on the occupation of all resident employees in Switzerland.[14]

We build a relative measure based on the values from Equation 2.3 for cross-border commuters (*Sh commuters*$_j$) and resident workers (*Sh residents*$_j$).

$$\text{Relative skill supply}_j = \frac{\text{Sh commuters}_j}{\text{Sh residents}_j}, j \in [1, 22] \tag{2.4}$$

[14]We focus on occupations held by resident workers living in the border region to control for potential differences in the industrial structure of places where cross-border commuters and resident employees work.

The measure *Relative skill supply*$_j$ indicates how the highly educated commuters are allocated across study fields j relative to the workers living in the country. In other words, a higher value of the measure implies that commuters are relatively more likely to have received training in this specific field than resident workers. In Table 2.7 we present for each study field the skill supply of commuters relative to that of resident workers. The least affected fields, those with the lowest ratio, are listed first in column (2) and the most affected fields come last. Education, languages and law have the lowest ratio and are non-STEM fields. Architecture and construction, information and communication technologies and forestry have the highest ratio and are STEM fields. Comparing columns (1) and (2) in Table 2.7 makes clear that there is a strong link between affected and STEM fields.

Table 2.7: Cross-border commuter shock by field of study

	STEM field (1)	Skill supply of commuters rel. to residents (2)
Education	0	0.529
Law	0	0.562
Languages	0	0.568
Welfare	0	0.577
Journalism and information	0	0.595
Personal services	0	0.616
Humanities (except languages)	0	0.687
Social and behavioral sciences	0	0.698
Business and administration	0	0.853
Health	0	0.940
Veterinary	0	0.955
Arts	0	1.068
Mathematics and statistics	1	1.391
Biological and related sciences	1	1.432
Agriculture	1	1.575
Manufacturing and processing	1	1.652
Environment	1	1.683
Physical sciences	1	1.687
Forestry	1	1.991
Engineering and engineering trades	1	2.073
Architecture and construction	1	2.478
Information and communication technologies (ICT)	1	2.972

Note: Two-digit ISCED study fields considered. STEM (science, technology, engineering, mathematics) and non-STEM fields are distinguished. Column (2) shows the ratio of the share of commuters trained in a field relative to the share of residents trained in the same field according to Equation 2.4. Source: EHA (2003–2017), FSO (1999, 2000).

2.5.2 Demand for Affected versus Non-Affected Fields

The literature on the occupation choice of foreign-born workers suggests that STEM skills are more transferable across countries than non-STEM skills (Hunt and Gauthier-Loiselle, 2010). STEM fields build analytical and quantitative abilities compared to non-STEM skills that often require institutional or cultural

knowledge or a high level of language skills. In their study, Hanson and Slaughter (2016) similarly observe in the US context that high-skilled immigrants are more likely to be employed in STEM than in non-STEM professions. The authors conjecture that these trends are explained by a higher quality of STEM training offered abroad or a lack of cultural and communication skills, which likely result in an absolute disadvantage for non-natives. In our context, we rule out the first argument because the largest Swiss institutions providing tertiary level STEM education and research are world leaders.[15] We hence document that STEM skills persist as more transferable even among foreign workers who have languge proficiency and are culturally similar. Hanson and Liu (2016) find that occupational comparative advantages specific to the country of origin persist even after adapting to the culture and language of the receiving country.

We divide the study fields into those with a value of the variable Relative skill supply$_j$ above and below one, where the former are referred to as "affected" and the latter as "non-affected". If natives try to avoid direct competition with foreign workers in the labor market, demand for less affected skills would increase in the treated regions. In the analysis of enrolment by field of study we find evidence that this is indeed the case. Table 2.8 shows a rise in enrolment in non-affected subjects in the post-treatment period of 2 percentage points, which is a sizeable growth of 7.8% relative to the mean of the outcome. The rise in the group of non-STEM fields is also statistically significant and of similar magnitude.[16] Figure 2.5 shows that the timing of the effects is in line with the implementation of the free movement reform. In contrast to Ransom and Winters (2020) who estimate crowding-out effects from STEM fields in regions with more foreign workers, we find no such evidence.

The higher enrolment in non-affected fields shows that natives in treated regions specialise more often in skills that are required in occupations less exposed to competition from foreign-born workers and that are less transferable across countries. This pattern is observed despite increasing wages paid in STEM occupations, as Table 2.9 shows. The returns to STEM degrees rise in the transition and post-treatment period, while returns to non-STEM degrees do not change significantly.

[15] In the academic year 2019/2020, the ETH ranked 6th and the EPFL 18th out of 1,001 in the QS World University Ranking. In the same year, the ETH ranked 13th and the EPFL 38th out of 1,001 in the THE World University Ranking.

[16] The range of the coefficient for enrolment in non-affected subjects is between 1.4 and 2.6 percentage points when leaving out one of the sixteen labor market regions one by one. The range of the coefficient for enrolment in non-STEM subjects is between 1.6 and 2.8 percentage points. In both exercises, the coefficients are statistically significant at least at the 5% level with two exceptions.

Table 2.8: Enrolment by type of study field

	Outcome: share of enrolled native first-year students			
	Affected	Non-affected	STEM	Non-STEM
	(1)	(2)	(3)	(4)
30min * 2002-2007	-0.000	0.003	-0.001	0.004
	(0.003)	(0.006)	(0.003)	(0.006)
30min * after 2008	0.003	0.020	0.002	0.021
	(0.003)	(0.009)	(0.003)	(0.009)
Mean outcome	0.117	0.254	0.104	0.267
Sd outcome	0.032	0.093	0.029	0.098
Commuting zones	106	106	106	106
within 30 min	35	35	35	35
N	2226	2226	2226	2226

Note: This table shows difference-in-differences estimates using annual data at the commuting zone level for the period 1997–2017. Affected fields are those with a supply shock measure above one as shown in Table 2.8. STEM fields include science, technology, engineering, mathematics. Outcome is the share of native first-year students in birth cohort. Denominator is fixed at first year. Observations are weighed by cohort size in first year. Standard errors in parentheses are clustered at the commuting zone level. Source: SHIS-studex.

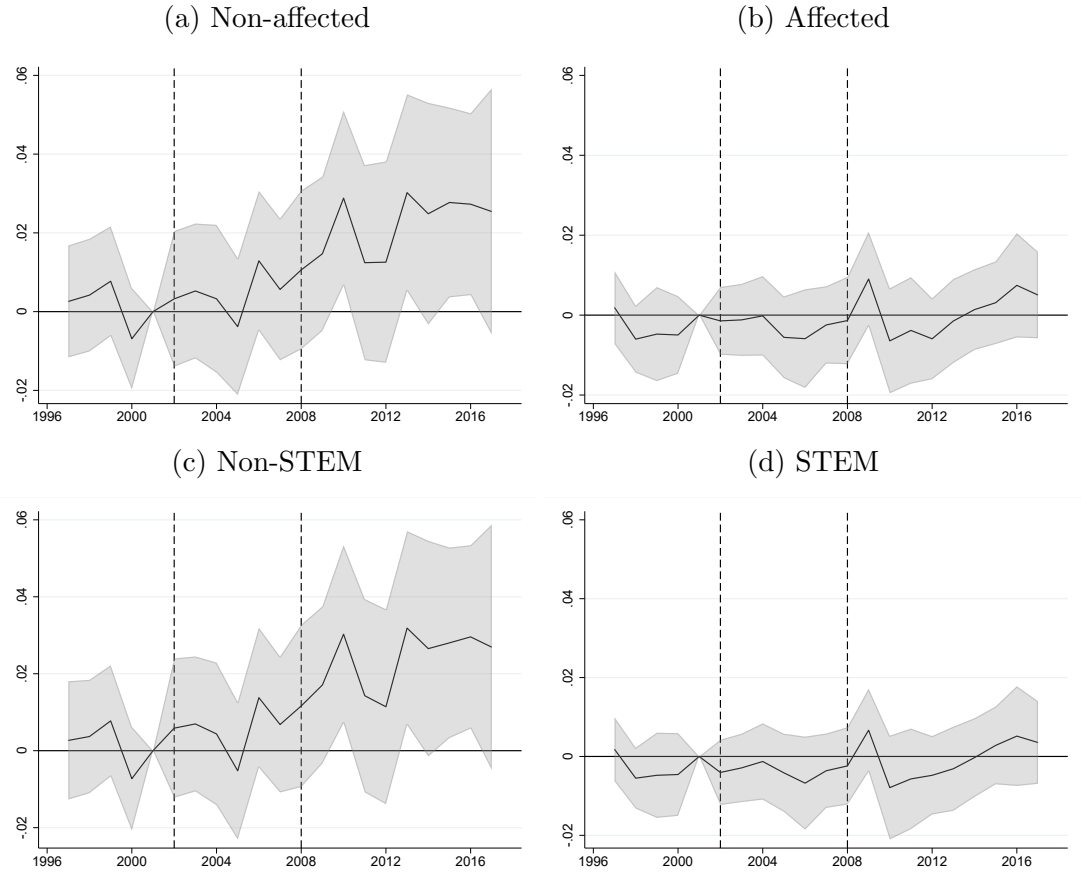
Table 2.9: Wages of tertiary educated by study field

	Outcome: ln **gross hourly wage rate of natives**	
	STEM	Non-STEM
	(1)	(2)
30min * 2002-2007	0.040	0.012
	(0.018)	(0.021)
30min * after 2008	0.038	0.027
	(0.022)	(0.028)
Mean outcome	3.909	4.027
Sd outcome	0.086	0.107
Commuting zones	94	105
within 30 min	34	35
N	1001	1144

Note: This table shows difference-in-differences estimates using biennial data at the commuting zone level for the period 1996–2016. STEM fields include science, technology, engineering, mathematics. Outcome is the mean natural log of gross hourly wage of tertiary educated natives in a region. Observations are weighed by number of tertiary educated native employees in first year. Standard errors in parentheses are clustered at the commuting zone level. Source: SESS.

In the literature on study field choice there is an established link between foreign students in a cohort and the choice of natives. Previous studies have shown that a higher share of foreign students reduces natives' demand for STEM subjects (Anelli et al., 2018; Orrenius and Zavodny, 2015). If tertiary institutions close to the border receive an increasing share of foreign students over time, this may alternatively drive natives into non-STEM fields – a result similar to what we observe. In our setting, first-year enrolment and study field choice takes place before the beginning of the studies and hence before knowing the composition of the cohort. This sequential timing mitigates the likelihood that our results are

Figure 2.5: Enrolment by type of study field



Note: This figure shows difference-in-differences estimates using annual data at the commuting zone level for the period 1997–2017. The vertical lines indicate the beginning of the transition period (2002) and of the post-treatment period (2008). Affected fields are those with a supply shock measure above one as shown in Table 2.8. STEM fields include science, technology, engineering, mathematics. Outcome is the share of native first-year students enrolled in specific group of fields relative to birth cohort. Denominator is fixed at first year. Observations are weighed by cohort size in first year. Standard errors are clustered at the commuting zone level, 95% confidence intervals shown. Source: SHIS-studex.

affected by foreign students.

2.6 Conclusion

We examine the impact of the introduction of free movement of workers on native demand for tertiary education. We find an increase in tertiary enrolment and link it to lower opportunity costs of studying and higher labor market rewards to tertiary degrees. Students with a vocational training at the upper-secondary level are more sensitive to changes in labor market returns as their higher demand for degrees at universities of applied sciences shows. The education system in the Swiss context grants access to tertiary degrees to individuals with a vocational and general background at the upper-secondary level. At the tertiary level, they usually enrol at different institutions with a focus on specific or general skills,

respectively. This institutional feature contributes to a highly educated labor force with a diverse skills set. It also gives vocationally trained an important margin of adjustment as shown in this study.

We also study how the reform affects study field choice and find no evidence that it is impacted by changes in monetary returns. We show that affected natives choose fields that build skills less transferable across countries compared to the STEM skills that the foreign workers typically bring to the labor market. Initial occupational specialization of high-skilled native and foreign workers could, thus, become more pronounced over time. Changes towards a more restrictive migration policy or deteriorating relative economic conditions in the host country can lead to a sudden outflow of foreigners. Such reversals could create a shortage of skills that foreign workers were previously supplying. Since skill acquisition is typically a long-term process, these findings should be taken into account when considering changes to immigration policies.

3 How to Improve Tax Compliance? Evidence from Population-wide Experiments in Belgium

with Jan-Emmanuel De Neve, Clément Imbert, Johannes Spinnewijn and Maarten Luts

3.1 Introduction

Tax compliance sits at the heart of the healthy functioning of societies. It is therefore of little surprise that gaining a robust understanding of the drivers of tax compliance is an important topic in the economics literature. Tax compliance involves both the truthful reporting of taxable income and the timely payment of tax dues. The growth in third-party reporting of income has limited the ability to misreport income (see Kleven et al. (2011, 2016); Jensen (2019)).[1] Tax administrations, however, continue to devote considerable resources to the collection of taxes. In the United States the annual cost of non-compliance with individual income taxes due to nonfiling, underreporting, and underpayment is estimated to total about $319 billion (Internal Revenue Service, 2016). Closing the "tax gap" is a key objective for governments around the world, and requires to know the drivers of tax compliance and the cost effectiveness of further interventions (OECD, 2010; HM Revenue & Customs, 2018).

The classic work by Allingham and Sandmo (1972) provided a work-horse model for understanding tax compliance through pecuniary incentives that deter non-compliance. Since then, a large body of research has stressed the role of non-pecuniary motives more broadly (e.g., Kirchler (2007); Luttmer and Singhal (2014); Besley et al. (2019)), often referred to as *tax morale*. There is now scattered evidence for these different drivers of tax compliance to be important across a variety of settings (see Slemrod (2018)), but several questions remain unanswered. In particular, while information frictions and complexity are shown to be important in related contexts (e.g., Bhargava and Manoli (2015); Cox et al. (2018)), their role in the context of tax compliance is less understood.

This paper studies the simplification of the communication by the tax authority and compares its impact on tax compliance to, on the one hand, the use of

[1] Recent empirical work investigates the misreporting of foreign income in developing countries (e.g., Alstadsæter et al. (2018)) and of taxable income in developing countries (e.g., Pomeranz (2015); Naritomi (2018)) where paper trails are missing or the enforcement capacity falls short.

deterrence and tax morale nudges and, on the other hand, the use of standard enforcement measures. We study compliance effects throughout the tax process – including the timing of tax filing, the reporting of taxable income, and the payment of taxes – for all individuals subject to personal income taxation in Belgium. We compare the potential drivers of tax compliance in the same context and put them on equal footing by varying the content of the tax letters sent by the Belgian tax authority (Federal Public Service Finance, FPS Finance). In total, we ran five population-wide natural field experiments in collaboration with the FPS Finance over the course of three fiscal years, 2014-2016. This comprehensive approach allows us to replicate findings at different stages of the tax process and across fiscal years, and to estimate longer-term, repetition and interaction effects.

The standard communication from the tax administration to taxpayers consists of a request to file a tax return and a request to pay taxes. Follow-up correspondence takes place in the event of taxpayers being late either in filing their tax return or in paying their tax dues. In order to estimate the impact of simplification and compare it to the use of deterrence or the appeal to tax morale, we leverage the different phases of communication and simultaneously test a variety of treatments. The simplification treatments shorten the length of the letters, reduce the information overload and highlight the action-relevant information to the taxpayer. The deterrence treatments add a message to the simplified letter that makes the financial penalties explicit and/or highlight the enforcement actions in case of non-compliance. The tax morale treatments add a message that highlights the public good value of tax expenditures and/or the social norms attached to filing and paying taxes on time.

Our experiments provide precise and remarkably consistent results across the tax process and the respective samples of taxpayers addressed. We find the largest compliance effects for the simplification treatments. Simplified tax filing reminders increase subsequent tax filing by 8% (relative to the baseline reminder). Simplifying the tax letter sent to all taxpayers with a positive tax bill increases timely payment by 0.7%.[2] For the late tax payers, the simplified reminder increases subsequent tax payment by as much as 23% (relative to the baseline reminder). Reducing information overload and emphasizing action-relevant information seem particularly effective in increasing compliance. We find that adding tax deterrence messages further increases tax compliance, with the average effect often being comparable in magnitude to the effect of simplification.

[2]Despite tax withholding one out of three taxpayers has a positive outstanding balance on their tax bill, adding up to a total of 3.8 billion euros in 2016 (about 10 percent of personal income taxes).

Tax payers are successfully induced to comply by making potential penalties and their enforcement explicit, and by the encouragement to pay or file immediately to avoid these penalties. In contrast, treatments that seek to improve tax morale obtain no compliance effects and sometimes even backfire. The ineffectiveness of tax morale messages is replicated across all treatments arms, which include messages that invoke social norms and/or emphasize the social value of public expenditures. For the latter, we also experiment with a pop-up pie chart of government expenditures for online tax filers and find that it does not affect reported taxable income, but neither does it affect the perceived importance of honesty as measured in an endline survey. While the survey shows that the treatment does increase taxpayers' knowledge and appreciation of public services, this seems insufficient to increase tax compliance.

More timely tax payments do not necessarily translate into greater tax revenues. In particular, we study the full dynamics of the treatment effects on late payers, and find that they diminish over time as the tax administration takes further enforcement measures (including imposing garnishments and sending bailiffs) to eventually reach close to full compliance. The simplification treatment effects at the end of the tax cycle are 1.0pp, which is ten times smaller than their effect at the payment deadline. Still, the cost savings on follow-up enforcement imply a large return to the simplification treatment. We exploit an enforcement discontinuity, combined with our experimental variation, to disentangle their respective effects. We estimate that the simplification treatment would have increased compliance by 5.2pp in the absence of enforcement actions, and that it is six times more cost-effective than standard enforcement.

Our empirical setting thus allows us to push the frontier on the evaluation of letter treatments by comparing their compliance effects to standard enforcement actions. While nudges are by definition low-cost interventions, knowing how they compare to the standard policy levers that they complement has been a key challenge (Benartzi et al., 2017). The enforcement discontinuity allows us to compare the causal impact of regular enforcement interventions and the experimental letter treatments for the exact same people (i.e., late taxpayers around the enforcement threshold). Projected on the sample of late taxpayers, whose tax liability was about €434 million, a back-of-the envelope calculation tells us that the simplification treatment for this experiment alone could have increased tax collection by €17.5 million, or alternatively, amounted to savings on enforcement costs worth €5.4 million. In comparison, the costs of the nudge intervention were trivial (€79,511).

Our experimental design also allows us to tackle a second important concern for the evaluation of letter interventions and nudge interventions more generally, which is whether the gains are long-lived (Allcott and Rogers, 2014; Cronqvist et al., 2018). To that purpose, we repeated the experiment on the late taxpayers in two consecutive years. We first find that there are no diminishing marginal returns to repeating the treatment in that recidivists are equally responsive to a simplified letter independent of the letter type they received in the previous year. Moreover, we find that the effects extend to the following fiscal year: late payers are less likely to be late again in the next year after having received a simplified reminder letter in the first year, but this effect is offset if they received a tax morale treatment as well. These effects become smaller, and statistically insignificant two years after the intervention.[3]

The particular features of our experimental setting help advancing the growing literature on randomized controlled tax trials and the evaluation of nudge-type interventions. More generally, our paper aims to contribute to the rich literature that studies the drivers of tax compliance (see Slemrod (2018)).[4] The first contribution of our paper is to focus on the role of complexity as a behavioral driver of tax compliance. While we do not address the complexity of the tax schedule itself (e.g., Chetty and Saez (2013), Abeler and Jäger (2015), Aghion et al. (2017)), our paper does shed new light on how simplifying communication can help to overcome information frictions and/or hassle costs associated with the process of filing and paying taxes (see e.g., Slemrod et al. (2001); Kleven and Kopczuk (2011); Hoopes et al. (2015); Dwenger et al. (2016); Benzarti (2017)). Relatedly, but in another context, Bhargava and Manoli (2015) identify barriers to the take-up of EITC benefits due to information complexity – with the mere simplification of the mailing leading to a significant increase in take-up. Second, we do not only show that simplifying the communication of the tax administration has a substantial effect on tax compliance, but also that this effect can outweigh the effects of nudges related to deterrence and to tax morale. Our study compares these various drivers of tax compliance in the same way, in the same setting, and on the same sample, which ensures comparability. This is particularly valuable as the results in the literature on tax morale are mixed. A number of experiments

[3]These findings extend on Brockmeyer et al. (2019), who find sustained effects from a deterrence message on firms' tax compliance in Costa Rica. These findings differ from Guyton et al. (2016), who find no long-term effects and positive returns from repeating reminders in claiming EITC.

[4]On the role of enforcement and deterrence, see reviews by Andreoni et al. (1998) and Slemrod and Yitzhaki (2002). An example of an RCT changing audit probabilities is Kleven et al. (2011). An example of an RCT changing the penalty information is Cranor et al. (2018). On the psychological, cultural, social, and normative factors underlying tax compliance, see Torgler (2007); Alm (2012); Luttmer and Singhal (2014) .

have found positive impacts from invoking social norms on tax compliance (e.g., Del Carpio (2014); Bott et al. (2017); Hallsworth et al. (2017); Perez-Truglia and Troiano (2018)), while several other experiments testing normative appeals have found null or even negative results (e.g., Blumenthal et al. (2001); Fellner et al. (2013); John and Blume (2018); Cranor et al. (2018)).[5] Third, we ran five population-wide natural field experiments that changed the communication between the tax authority and tax payers, which allows us to test the effects of the interventions at scale, at all stages of the tax process and for different subsets of the tax payer population, thus strengthening the internal validity of our design.[6]

The paper proceeds as follows. Section 2 presents a simple model of tax compliance and characterizes the cost-effectiveness of different interventions. Section 3 describes the context and empirical setting. Section 4 discusses the main experimental results, presents the dynamics and sheds some light on mechanisms. Section 5 analyzes the regression-discontinuity in enforcement, compares the cost-effectiveness of simplification with traditional enforcement and studies its long-term effects. Section 6 concludes.

3.2 Model

We consider a stylized model of tax compliance, revisiting the model of criminal behavior in Becker (1968) and its adaptation to tax evasion by Allingham and Sandmo (1972). A taxpayer decides whether to comply with their tax duties, which include the accurate reporting of their taxable income y and the timely filing and payment of taxes dues $\tau(y)$. We model tax compliance behavior as an action $\tilde{y} \in [0, y]$, which solves

$$\min_{\tilde{y} \in [0,y]} T(\tilde{y}) + \Phi_{non-compliance}(y - \tilde{y}) + \Phi_{morale}(y - \tilde{y}) + \Phi_{compliance}(\tilde{y}),$$

with $T' \geq 0$ and $\Phi_j' \geq 0$. The first and most natural cost from complying is the loss of resources from paying taxes, $T(\tilde{y})$. However, by complying the taxpayer can avoid follow-up costs enforced by the tax authority, captured by $\Phi_{non-compliance}(y - \tilde{y})$. This is the central trade-off in the deterrence framework

[5] For example, Hallsworth et al. (2017) find that social norms and public services messages in official reminder letters increased payment rates for overdue tax in the UK. In contrast, Cranor et al. (2018) find that invoking social norms has no compliance effects on late tax payers in Colorado, while making the penalty explicit does. Another recent example is Perez-Truglia and Troiano (2018), who find that shaming tax payers by making their non-compliance public increases compliance. However, they find no effects from providing information on others' non-compliance.

[6] We also test different variations of similar treatments and study heterogeneous treatment effects with causal forests (Wager and Athey, 2018), which helps to establish robustness and uncover underlying mechanisms.

by Allingham and Sandmo (1972), where the tax authority increases the costs of non-compliance by increasing penalties for non-compliant behavior and the probability of actual enforcement.[7] In addition to the resource costs, taxpayers may also face an intrinsic cost of non-compliance given their *tax morale*, captured by $\Phi_{morale}(y - \tilde{y})$. This cost may depend on the perceived fairness of the tax system, the taxpayer's valuation of the government's use of the tax revenues, social norms determined by the compliance behavior of other tax payers, etc. Finally, we also allow for a direct cost of compliance $\Phi_{compliance}(\tilde{y})$, which can capture the hassle cost of filing and paying taxes, the attention needed in order to take the appropriate action, etc.

To induce compliant behavior, the tax authority needs to ensure that the cost of compliance is exceeded by its return. Assuming linear cost functions ($H(x) = h \times x$), this can be represented by

$$t + \phi_{compliance} \leq \phi_{non-compliance} + \phi_{morale}.$$

The tax authority has a set of instruments available that can affect the vector of cost parameters ϕ determining the taxpayer's compliance $\tilde{y}(\phi)$. This includes standard enforcement interventions (which affect compliance through $\phi_{non-compliance}$), but also the letter interventions that we consider below. We categorize our interventions as affecting $\phi_{compliance}$ through simplifying/improving the letter design, $\phi_{non-compliance}$ through making enforcement and penalties explicit, and ϕ_{morale} by invoking tax morale.

The optimal mix of instruments will depend on their cost effectiveness, determined by their impact on tax revenues $\partial T/\partial \phi_j$ and their resource cost to the tax authority $\partial C/\partial \phi_j$. Leaving aside other considerations, the tax authority should equalize the marginal cost of raising an extra euro of revenue $\frac{\partial C/\partial \phi_j}{\partial T/\partial \phi_j}$, as shown by Keen and Slemrod (2017). In practice, especially in the case of payment recovery, the tax authority may aim to reach near full compliance $\tilde{y}(\phi) \approx y$ and rely on stronger enforcement to recover the remaining taxes due. In that case, the return to alternative interventions is not the increase in tax revenues, but the costs savings on the standard enforcement measures. The relative cost-effectiveness of the alternative intervention can then be written as

$$\frac{\frac{\partial C}{\partial \phi_{non-compliance}}}{\frac{\partial C}{\partial \phi_j}} \times \frac{d\phi_{non-compliance}}{d\phi_j}|_{\tilde{y}(\phi)=y}.$$

[7] Note that the cost $\Phi_{non-compliance}(y - \tilde{y})$ can also include the resources taxpayers expend to camouflage non-compliance (see Slemrod (2018)).

This is exactly the metric we will calculate after having estimated the compliance effects and costs of the letter interventions and standard enforcement.

3.3 Context and Design

This section presents the five experiments we study and describes the experimental samples. We also provide some background on the tax filing and payment cycle for personal income taxation in Belgium.

3.3.1 Tax Process

In Belgium the tax-to-GDP ratio was 44.6% in 2017, which is above the OECD average of 34.2%. We focus on individual income tax, which is the largest source of tax revenues in Belgium. In the fiscal year 2016, individual income tax raised 27.7% of overall tax revenues from 7.1 million taxpayers. Income taxes are collected solely at the federal level. There is a personal tax-free allowance which stood at 7,130 EUR and marginal taxes rise from 25 to 50%.[8] Fiscal years run from January 1st to December 31st, and the tax cycle starts in July of the year after the fiscal year in which the income has been earned. There are four main steps in the annual personal income tax cycle, as shown in Figure 3.1a: tax filing, filing reminders, tax payment and payment reminders. We vary the correspondence between the tax administration and taxpayer at each of these steps.

Tax filing (TF): Taxpayers can file their taxes on paper or online, either by themselves or with the help of an accountant or a tax official.[9] The online portal called "Tax-on-Web" is increasingly popular and in 2017 it was used by 3.8 million taxpayers, of which 1.7 million submitted their declarations individually. The remainder filed with the help of an accountant or a government official.

Filing reminders (TFR): Figure 3.1b depicts what happens when taxpayers miss the filing deadline. Filers who have not submitted by the deadline are sent a filing reminder letter, and given 14 days to file. If a taxpayer has still not filed seven days after this second deadline, the tax administration uses its own estimates to compute their tax liability. In the fiscal year 2016, about 170,000

[8] In comparison, in the US, the tax-to-GDP ratio is lower (27.1%) and income taxes are more important as a share of tax revenues (38.6%). Federal marginal tax rates are lower (10 to 37%), but lower levels of government levy additional taxes.

[9] Not all taxpayers need to file. About a third of taxpayers (2.2 million in the fiscal year 2016) receive pre-filled tax returns with no further action required.

taxpayers had not filed by the deadline, which represents about 3.5% of taxpayers who were expected to file.

Tax payment (TP): A majority of taxpayers are taxed at the source if they are employed or pre-pay their taxes based on estimates of their tax liability if they are self-employed. A significant share of taxpayers also have taxable income below the exemption threshold and thus pay no income taxes. As a result, less than a third of taxpayers (1.9 million in the fiscal year 2016) receives a tax bill with a positive payable balance, which they need to pay within the next two months. The majority of such cases can be explained by insufficient withholding at the source in situations that made it difficult to calculate the exact tax liability (e.g. tax payers who hold several jobs, students who work part-time, etc.). Total taxes due at that stage are 3.8 billion euros.

Payment reminders (TPR): Figure 3.1c depicts what happens when taxpayers miss the payment deadline. Taxpayers who have not paid two months after receipt of the tax bill are sent a payment reminder. Taxpayers who still do not comply are then exposed to further enforcement actions, which start after 14 days. In the fiscal year 2016, about 220,000 taxpayers had still not paid 14 days after the deadline, and owed a total of 0.8 billion euros, which represents 12% of taxpayers who received a positive tax bill, and 21% of taxes they owed.

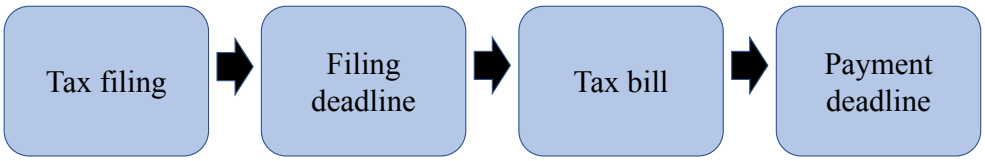
3.3.2 Experiments

We report on a total of five experiments: one on tax filing (TF), one on tax filing reminders (TFR), one on tax payment (TP) and two on tax payment reminders (TPR). The experiments spanned the three fiscal years (FY) from FY2014 to FY2016. The experiments involve various randomly assigned treatments that we categorize in three groups: simplification, deterrence and tax morale.

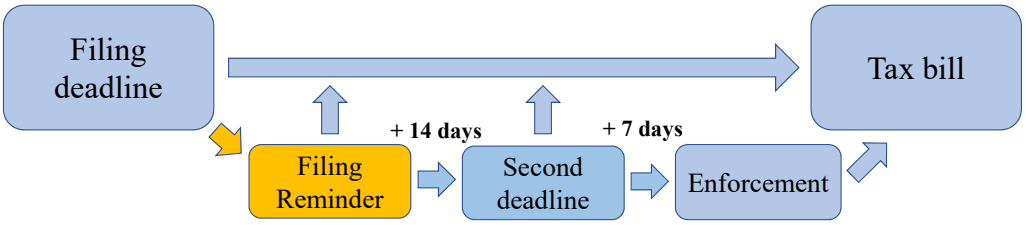
In four experiments out of five, the treatment involved simplifying the letter to communicate more clearly what the tax administration expected from taxpayers. Simplification included shortening the letter while retaining the action-relevant information. To attract the attention of the reader, important information was highlighted in color and/or placed in boxes. The simplified letters were also personalized, i.e., it was addressed to the taxpayer using his/her name.[10] As we

[10]Only for the TP experiment, we have within-experiment variation in the design of the simplified letter as the non-personalized address is used for a random subgroup.

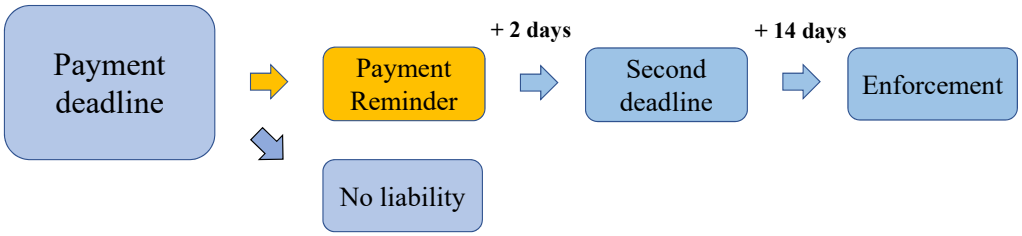
Figure 3.1: Tax process



(a) Filing and payment



(b) Filing reminder process



(c) Payment reminder process

discuss below, the exact design of the simplified letter varies across experiments as does the design of the old letter.

The experiments also tested the effect of deterrence and tax morale through the addition of short messages in the simplified letter. The deterrence messages aimed at making the consequences of non-compliance explicit, by stating fines and tax increases and/or by mentioning follow-up enforcement. We also tested messages that encouraged immediate action to avoid the fines. The tax morale messages, on the other hand, aimed at raising compliance by increasing the desire of taxpayers to comply with social norms or to reciprocate for public goods provision. Appendix Table C1 lists all the deterrence and tax morale messages used (translated in English).

TP Experiment: The Tax Payment experiment modified the tax bill sent to taxpayers with a positive liability: the experiment was carried out between November 2017 and May 2018 with 1,216,317 taxpayers (fiscal year 2016). All treated taxpayers received a simplified letter, only keeping action-relevant information and improving the overall outline. For a subset of treated individuals, the letter included either deterrence messages or tax morale messages (see Panel A of Appendix Table C1). For this experiment, outcomes include the probability of making a payment following letter receipt (extensive margin response), and the fraction paid conditional on a payment having been made (intensive margin). As baseline outcome, we use the probability of payment within 60 days after the letter was sent: 60 days is the deadline given to taxpayers to pay their outstanding debt.

TPR Experiments: The Payment Reminder experiments were conducted with taxpayers who were late in paying their tax: 229,751 taxpayers in 2015/16 (FY2014) and 188,180 taxpayers in 2016/17 (FY2015).[11] The treatment group received a simplified reminder letter, in which the outstanding tax liability and the deadline were highlighted and other information shortened. Again, for different subsets of the treatment group, the letter also included deterrence and tax morale messages (see Panel B of Appendix Table C1). The baseline outcome we consider is now the probability of payment within 14 and 180 days after reminder receipt: 14 days corresponds to the time at which enforcement actions

[11] In both trials, German speaking taxpayers, taxpayers who had raised objections to the outstanding amount they owed and taxpayers for whom the government did not have a name were not included in the randomization and received an old letter. Only debts related to the current fiscal year and letters that are first means of communication with the taxpayer (no updates on balances owed) are included in the analysis.

begin. To validate the results and to test the effect of repeated treatments, the TPR experiment was conducted in two consecutive years.

TF Experiment: The Tax Filing experiment was conducted in 2017 (FY2016) with 1.5 million online tax filers.[12] The tax filers were shown a pop-up pie chart either before (treatment) or after (control group) they filed their taxes. The pie chart presented the breakdown of government spending by categories.[13] The chart was accompanied by a sentence highlighting that these public services were funded by taxes.[14] We consider this as a similar treatment to the tax morale message in the other experiments. For this experiment, outcomes come from two sources: administrative data on tax compliance and answers to an online survey to which all online filers were invited. Due to confidentiality concerns, the administration did not provide individual information but only average outcomes (or taxpayers characteristics) within a gender-age cell. The main compliance outcome is reported taxable income. Other outcomes are tax liability, self-employed profits and expenses, expenses of salaried workers and general expenses. These are also based on declared values. Survey data is available for those who agreed to answer the questionnaire, which gauges taxpayers' knowledge and agreement with the way tax revenue is spent, and their evaluation of public services and the tax system more generally. [15]

TFR Experiment: The Filing Reminders experiment was conducted with 148,925 taxpayers who were late in filing their tax returns in 2016 (FY2015). The treatment group received a simplified letter, which emphasized the new filing deadline. A subset of the treatment group received a letter which included deterrence messages (see Panel D of Appendix Table C1).[16] For these experiments, the baseline outcome is the probability of filing within 21 days after letter

[12]This excludes taxpayers who used an accountant or tax officer to submit their taxes via the online portal. Our dataset covers taxpayers who submitted their tax returns before mid-August 2017.

[13]The tax administration also provided a pie chart of government expenditures by region, which was available when scrolling down.

[14]For some randomly selected sub-groups, the administration added at the very bottom of the pop-up an additional sentence that either added a public goods message, mentioned penalties in general terms, or appealed to social norms in general terms (see Panel C of Appendix Table C1). We do not find any differential effect of this second sentence and pool all treatment groups in the analysis.

[15]All outcome variables were pre-specified in the Pre-analysis Plan (AEARCTR-0002196).

[16]In the previous year (FY2014), the administration carried out a separate experiment on filing reminders, in which it included tax morale messages without simplifying the letter first. We managed to collect data from this experiment and found no effect of the treatment. Results are not reported here.

receipt: 21 days is the time at which the tax administration begins to calculate the tax liability based on income estimates.

3.3.3 Randomization Design

The allocation of taxpayers to the different treatment groups was done in two different ways. For the TPR, the TF and the TFR experiments, it was based on the last two digits of the national identity number, which are random (see Appendix Table C2). For the TP experiment, treatment allocation was based on the day of the month the taxpayer was born, which is also random and independent of the last digits of the national identity number (see Appendix Table C3). There are three things to note.

First, treatment allocations for the two tax payment reminder experiments (TPR 2014 and the TPR 2015) were done in such a way that taxpayers of each treatment group in TPR 2014 had a similar probability to be assigned to each treatment group in TPR 2015. It follows that the two allocations are almost independent from each other, as in a cross-cutting randomization design.[17] Since there is significant overlap between 2014 and 2015 late payers (see Appendix Table C4), we have sufficient power to estimate the effect of the two treatments both separately and jointly, to identify the effect of repeated treatment.

Second, treatment allocations for the TPR 2014 (tax payment reminder) and TFR 2015 (tax filing reminder) experiments coincide partially, but not completely. A potential concern could be that treatment status in one experiment affects outcomes in a following experiment. Fortunately, the two experiments were done on different target populations, since the late payers of 2014 need not be late filers in 2015. Indeed, the overlap between the two populations is small: as Appendix Table C4 shows, only 6% of late payers for the fiscal year 2014 were also late filers for the fiscal year 2015. As a robustness check, we estimate the results of the TFR 2015 experiment controlling for the TPR 2014 treatment assignment and show that our results do not change.

Third, treatment allocation for the TF 2016 experiment again split the tax sample in two based on the two last digits of the national identity number, which made it partly, but not completely coincide with treatment allocations for the TFR and the TPR 2014 experiments. Unfortunately, to protect privacy the tax

[17]Since 97 digits had to be allocated to 9 treatment groups in TPR 2014 and 10 treatment groups in TPR 2015, the two allocations are independent up to seven digits (11, 22, 33, 64, 75, 86 and 97).

administration did not share individual identifiers for the TF 2016 experiment, which prevents us from measuring the exact overlap with the sample of the other two experiments, or controlling for assignment to previous treatments. However, since the sample of the TF experiment is much larger (1.5 million, against 150,000 for TFR and 230,000 for TPR 2014), the overlap is likely to be small.

3.3.4 Population comparison

As the five experiments take place at different stages of the tax process, they test the effect of simplification, deterrence and tax morale on different parts of the taxpayer population. Table 3.1 shows descriptive statistics on socio-demographic characteristics of the different experimental samples, as compared to the universe of Belgian taxpayers. The Belgian personal income taxpayer is on average 49 years old, in a couple in 35% of the time and has 0.4 children (column 1). By convention, in the case of households composed of individuals of both genders, only the gender of the woman is recorded, so that there are many more female than males (70%). 33% of the taxpayer population lives in Wallonia and 42% speak French. On average, they owe €570, but only 28% have a positive tax liability. Taxpayers in the TP experiment have a tax liability which is by definition positive, with an average of €2676. As column 2 shows, they are older, more likely to be in a couple and less likely to have children. In contrast, taxpayers in the TF experiment (column 4), who file online, are younger, and have more children. Taxpayers in the reminder experiments (TPR and TFR in columns 3 and 5) differ from the overall population in similar ways: they are more likely to be male, less likely to be in a couple, younger, more likely to speak French and to live in Wallonia. Taxpayers who are late in paying also have lower tax liability than the average (€1890). For late taxpayers, we were able to collect two additional covariates: taxable income and solvency score. The solvency score is the prediction by the tax administration of the probability that a taxpayer will not be able to pay their debts permanently, based on their tax returns in the previous year and their debt settlement history.

3.4 Experimental Results

This section first presents the main results of our experiments, then discusses the timing of the effect of the different interventions, and finally explores potential mechanisms.

Table 3.1: Summary Statistics of Control Variables

Experiment:	All taxpayers	Tax Payment	Payment Reminder	Tax Filing	Filing Reminder
	(1)	(2)	(3)	(4)	(5)
Demographics					
Male dummy	0.309	0.324	0.448	0.276	0.529
	(0.462)	(0.468)	(0.497)	(0.447)	(0.499)
Couple dummy	0.346	0.415	0.298	0.445	0.132
	(0.476)	(0.493)	(0.457)	(0.497)	(0.339)
Age	49.495	53.354	47.764	47.596	42.229
	(18.129)	(16.382)	(15.611)	(15.585)	(16.249)
Number of children	0.413	0.351	0.409	0.579	0.334
	(0.869)	(0.771)	(0.830)	(0.950)	(0.836)
Married dummy				0.476	
				(0.499)	
Widowed dummy				0.040	
				(0.196)	
Divorced dummy				0.156	
				(0.363)	
Region / Language					
Wallonia dummy	0.327	0.316	0.367	0.284	0.390
	(0.469)	(0.465)	(0.482)	(0.451)	(0.488)
Flanders dummy	0.570	0.596	0.525	0.637	0.390
	(0.495)	(0.491)	(0.499)	(0.481)	(0.488)
French dummy	0.421	0.386	0.473	0.357	0.592
	(0.494)	(0.487)	(0.499)	(0.479)	(0.491)
German dummy	0.006	0.011	-	0.003	0.007
	(0.076)	(0.104)		(0.051)	(0.084)
Other					
Amount owed	568.635	2676.205	1890.950		
	(7301.068)	(11869.230)	(4746.221)		
Income			33211.010		
			(28804.210)		
Solvency score			11.657		
			(4.674)		
N	6,689,779	1,216,317	229,751	942,571	148,925

Note: The table presents means and standard deviations (in parentheses) of control variables for different samples. In column 1 the sample is composed of all individual income taxpayers in FY2016. In column 2 it is the sample of the TP FY2016 experiment. In column 3 it is the sample of the TPR FY2014 experiment. In column 4 it is the sample of the TF FY2016 experiment. In column 5 it is the sample of the TFR FY2015 experiment. The base category for gender is female, for region Brussels, for language Flemish and for marital status single.

3.4.1 Baseline Results

To estimate the effect of simplification, deterrence and tax morale messages in each experiment, we take advantage of the randomization and simply regress compliance outcomes on treatment dummies and taxpayer controls. The estimating equation writes:

$$Y_i = \alpha + \beta_S S_i + \Sigma_j \beta_j T_i^j + \gamma \mathbf{X_i} + \varepsilon_i,$$

where Y_i is the relevant outcome for taxpayer i, S_i is a dummy variable equal to one for taxpayers who received a simplified letter, T_i^j are dummy variables equal to one for the different messages added to the simplified letter, and $\mathbf{X_i}$ is a vector of taxpayer characteristics.

The outcome variable Y_i we use for our baseline specification in the tax payment experiment is whether the tax liability is paid (in full or in part) before the deadline, which is 60 days after the letter receipt. For the reminder experiments, the outcome variable is whether taxes are filed or paid before the start of follow-up interventions (respectively after 21 and 14 days for the filing and payment experiments). We consider compliance at different time horizons and at the extensive vs. intensive margin later in this section. For the tax filing experiment, the compliance variable is different in nature, since we consider total reported taxable income. Table 3.1 presents the full list of controls $\mathbf{X_i}$. Controls include dummies for gender, couples, age, region, mother tongue, and number of children. For experiments in which letters were sent out in waves, controls also include dummies for each wave. We include additional controls for some experiments: dummies for quintiles of amount owed (TP and TPR experiments), quintiles of income and solvency score (TPR experiment), and marital status (TF experiment).

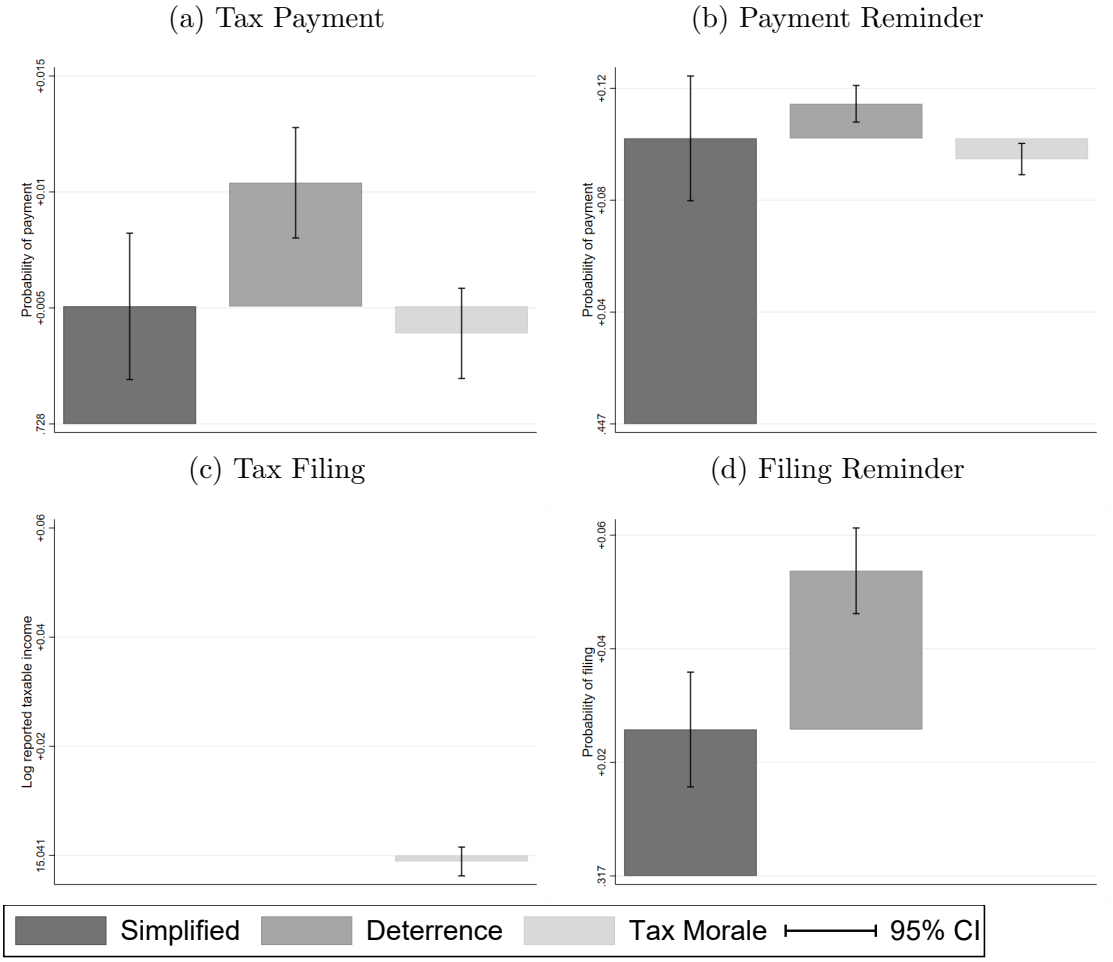
The coefficients of interest are β_S, which identifies the effect of simplification, and β_j, which identifies the effect of adding a deterrence or tax morale message.

Figure 3.2 presents our baseline estimates for the simplification, deterrence and tax morale treatment. The tax payment and tax filing experiments are in the top and bottom panels respectively. The experiments on the baseline sample of tax payers/filers are on the left, while reminder experiments for the late payers/filers are on the right. The figure conveys a very clear and strong pattern across the four experiments. In the three experiments in which communication with the taxpayer was simplified (TP, TPR and TFR), it had a positive and sizeable effect on tax compliance. In the same three experiments, the deterrence messages had an additional positive effect, which is significant and can be as large as the effect of simplification. Finally, in the three experiments in which the administration tried to increase tax morale (TP, TPR and TF), it had either no effect or even reduced compliance.[18]

The regression estimates are also presented in Table 3.2, which has the same structure as Figure 3.2. The top panel (Panel A) presents the results of the tax payment experiments. Column 1 shows that simplifying the tax bill had a positive effect on the probability of paying on time, increasing it by 0.5pp. Adding a deterrence message increased the probability of paying on time further, by 0.5pp. These effects are relatively small, but significant: the combined effect

[18]Another TFR experiment was run in 2014, but unlike the main 2015 experiment, only tax morale messages were used, and without simplifying the letter. These messages had a null or negative effect on the probability of filing before enforcement actions started. These results (not shown here) confirm that tax morale messages do not improve tax compliance.

Figure 3.2: Summary of the Main Results



Note: The figure presents treatment effect estimates from baseline specifications for the TP (Panel (a)), TPR FY2014 (Panel (b)), TF (Panel (c)) and TFR FY2015 (Panel (d)) experiments. The outcome is partial payment probability at 60 days (deadline) in Panel (a), and at 14 days (enforcement) in Panel (b). The outcome is reported taxable income in Panel (c) and filing probability at 21 days (enforcement) in Panel (d). Control variables are listed in Table 3.1, for exact estimates refer to Table 3.2. 95% confidence intervals based on robust standard errors are plotted. Standard errors are clustered by date of letter receipt in Panels (a) and (b).

of simplification and deterrence messages is 1.4% of the control mean (72.8%). The tax morale messages, however, had no additional effect on tax compliance. The effect of −0.1pp is sufficiently precisely estimated to rule out effects of a magnitude comparable to the simplification and deterrence treatment. Column 2 presents the results of the payment reminders experiment. The results are qualitatively similar. The effects of simplification and deterrence are again positive, but the former effect clearly dominates. That is, simplifying the reminder letters increased the probability of paying by 10pp (22.8% of the control mean), and deterrence messages had an additional positive effect of 1.2pp (2.7% of the control mean). Tax morale messages, however, had an opposite effect, slightly reducing

tax compliance (−0.7pp or 1.6% of the control mean). The bottom panel (Panel B) presents the results of the tax filing experiments, which are again very similar qualitatively. The tax morale treatment in the tax filing experiment (Panel B Column 1) had no effect on declared taxable income, with the null effect again being precisely estimated. The estimates in Column 2 of Panel B show that simplification and deterrence had a large positive effect on tax compliance among late filers. Those who received a simplified letter were 2.6pp more likely to file on time. This probability increased by an additional 2.8pp for those who received a simplified letter with a deterrence message, making them 17% more likely to file on time than the control group.[19]

3.4.2 Dynamic Effects

We have so far reported treatment effects at one point in time, at the deadline for the tax payment experiment and before the start of enforcement actions for the reminder experiments. Using the payment and filing history, we can estimate treatment effects at any time – measured in days – after treatment. Let $Y_{i,t}$ be the tax compliance outcome of individual i at time t. As before, S_i denotes a dummy variable equal to one for taxpayers who received a simplified letter, T_i^j are treatment dummies for the addition of deterrence and tax morale messages and $\mathbf{X_i}$ denotes a vector of controls. We estimate the following equation:

$$Y_{i,t} = \alpha_t + \beta_{S,t} S_i + \Sigma_j \beta_{j,t} T_i^j + \gamma \mathbf{X_i} + \epsilon_i.$$

For the TP experiment, t ranges from the receipt of the tax bill to 60 days after, corresponding to the deadline. For the TPR experiment, t ranges from the receipt of the letter to 180 days after. Note that the deadline is two days after, and that enforcement follow-up does not start until 14 days later. For the TFR experiment, t ranges from the receipt of the letter, which gives late filers 14 days to comply, to 60 days after, when the administration automatically files taxes for non-compliers.

Appendix Figure C1 displays the dynamics of tax compliance in the control group - the estimated α_t - for the three experiments. In the TP experiment, the proportion of taxpayers who paid in the control group increased slowly after receipt of the tax bill, and then sharply just before the deadline, so that 72%

[19] Appendix Table C5 presents the results of the filing reminder experiment controlling for the treatment assignment in the payment reminder experiment. Due to the partial overlap between the two experiments, the estimates are less precise, but the magnitude of the treatment effects is similar.

Table 3.2: Main Results

Panel A: Payment	Probability of some payment at 60 days (deadline)	Probability of some payment at 14 days (before enforcement)
	Tax Payment (1)	Payment Reminders (2)
Simplified (S)	0.005	0.102
	(0.001)	(0.010)
+ Deterrence	0.005	0.012
	(0.001)	(0.003)
+ Tax Morale	-0.001	-0.007
	(0.001)	(0.003)
P-values of tests:		
Simplified=Control	0.001	0.001
S+Deterrence=Simplified	0.001	0.001
S+Tax Morale=Simplified	0.167	0.083
Control mean	0.728	0.447
N	1,216,317	229,751
Panel B: Filing	Log pre-check taxable income	Probability of having filed at 21 days (before enforcement)
	Tax Filing (1)	Filing Reminders (2)
Simplified (S)		0.026
		(0.005)
+ Deterrence		0.028
		(0.004)
Tax Morale	-0.001	
	(0.001)	
P-values of tests:		
Simplified=Control		0.001
S+Deterrence=Simplified		0.001
Tax Morale=Control	0.413	
Control mean	15.041	0.317
N	942,571	148,925

Note: The table presents treatment effect estimates from baseline specifications in four separate experiments. Column 1 in Panel A presents the results of the TP experiment (taxpayers for the FY2016). Column 2 in Panel A presents the results of the TPR 2014 experiment (late taxpayers in the FY2014). Column 1 in Panel B presents the results of the TF experiment (online tax filers in the FY2016). Column 2 in Panel B presents the results of the TFR experiment (late tax filers in the FY2015). Control variables are listed in Table 3.1. Robust standard errors in parentheses, clustered by date of letter receipt in Panel A. p-values are adjusted for multiple hypothesis testing (List et al., 2016).

of taxpayers met the deadline. In the TPR experiment, only a minority of late payers (17%) met the renewed deadline, and less than half of them had paid before the beginning of enforcement actions. The pattern is similar in the TFR experiment: only 25% of late filers in the control group had filed by the renewed deadline and only 34% had filed before enforcement actions began.

Figure 3.3 presents the dynamics of the simplification treatment, $\beta_{S,t}$. Taxpayers who received a simplified tax bill were slightly more likely to pay in the first weeks after tax bill receipt, but the difference with the control group really widened in the last week before the deadline. For the late payers, who were given a tight

deadline, the simplified reminders had a strong and immediate effect on payment probability, which peaked around the time when enforcement actions started. As enforcement actions began, the control group caught up with treatment, so that the treatment effects decreased steeply, although they were still statistically significant at the end of the period. In the filing reminder experiment, the simplified reminders also had a strong and rapid effect on filing probability, which accelerated close to the deadline and peaked at the time at which enforcement actions started. Then, as income was automatically filed, the difference in manual filing remained constant between treatment and control. Taken together, these findings suggest that simplification made both the need to pay and the actual deadline more salient to taxpayers. For completeness, we also report on the dynamic effects of deterrence and tax morale messages, $\beta_{j,t}$, in Appendix Figure C2. Across the three experiments, the additional positive effect of deterrence messages, which emphasized the penalties associated with missing the deadline, were felt gradually, and peaked at the deadline. In the Payment Reminder experiment, the negative effect of tax morale messages lingered for about a month, even after enforcement actions begun.

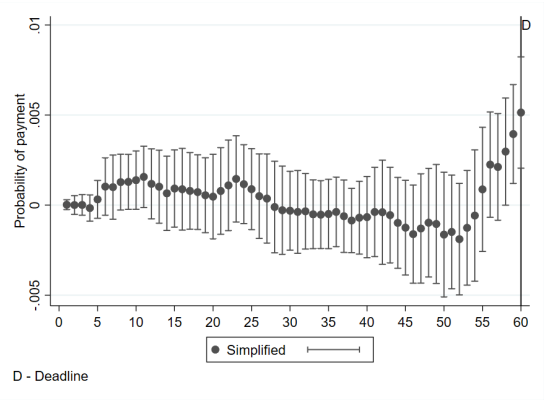
While our results show that the compliance effects peaked at the deadline or shortly after, they also clearly show that the effects diminished over time as enforcement actions begun. This is particularly striking in the TPR experiment. As Table 3.3 shows, compliance was 10pp higher in treatment than in control after 14 days (before enforcement), 6.9pp higher after 30 days, and less than 1pp higher after 180 days. Hence, the effect of simplification on taxes collected was in the end much smaller than the effect on compliance at 14 days would suggest. However, it declined in part because enforcement actions by the tax administration made the control group catch up with the treatment group. In Section 3.5, we will disentangle the compliance effect of the simplification treatment and the follow-up interventions.

3.4.3 Mechanisms

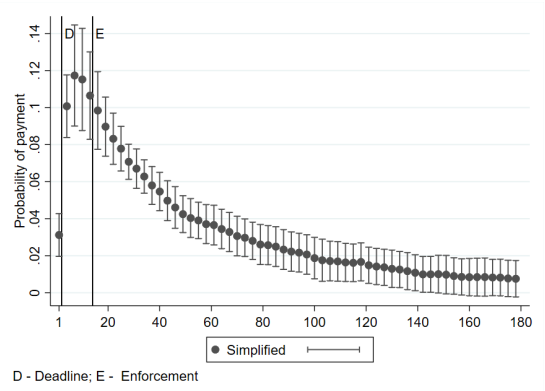
The relative impact of the simplification, deterrence and tax morale treatments is remarkably consistent across experiments implemented at different stages of the tax process, and on different populations. This section explores potential mechanisms underlying this robust pattern. We present treatment variations within each category, consider their impact on alternative outcome variables and present heterogeneous effects estimated with causal forests.

Figure 3.3: Dynamic Effects of Simplification

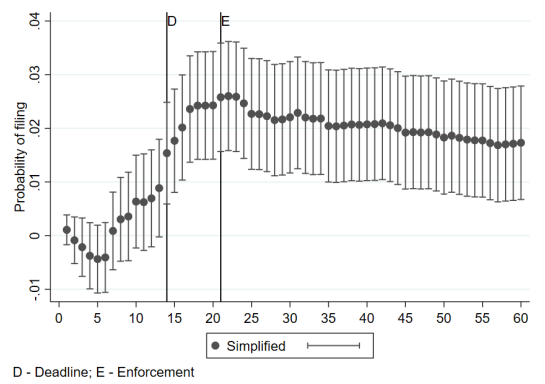
(a) Tax Payment



(b) Payment Reminder



(c) Filing Reminder



Note: The figure presents simplification treatment effect estimates by days since letter receipt for the TP (Panel (a)), TPR FY2014 (Panel (b)) and TFR FY2015 (Panel (c)) experiments. The outcome is partial payment probability in Panels (a) and (b), and filing probability in Panel (c). The vertical lines indicate the payment/filing deadline and/or the day enforcement actions start. Control variables are listed in Table 3.1. 95% confidence intervals based on robust standard errors plotted. Standard errors are clustered by date of letter receipt in Panels (a) and (b).

Simplification Our experiments show that simplifying the tax correspondence can have a substantial impact on compliance and highlighted the dynamic patterns of the compliance effects. We briefly compare the compliance effect across

Table 3.3: Dynamic Effects of Payment Reminders FY2014

	Probability of some payment			
	at 2 days (deadline)	at 14 days (before enforcement)	at 30 days (after enforcement)	at 180 days (after enforcement)
	(1)	(2)	(3)	(4)
Simplified	0.065	0.103	0.069	0.010
	(0.011)	(0.010)	(0.004)	(0.003)
P-values of tests:				
Simplified=Control	0.001	0.001	0.001	0.001
Control mean	0.166	0.447	0.598	0.845
N	229,751	229,751	229,751	229,751

Note: The table presents treatment effect estimates from the payment reminders experiment (TPR FY2014). Control variables are listed in Table 3.1. Standard errors in parentheses are clustered by date of letter receipt. p-values are adjusted for multiple hypothesis testing (List et al., 2016).

experiments and across slight treatment variations within one experiment.

To compare the magnitude of the effects of simplification across experiments, it is important to keep in mind that while the simplified letters look very similar, the quality of the old letters was different. In particular, in the tax payment experiment, the required actions were already grouped together and highlighted in the *old* letter, but they were made even more salient in the *new* letter. For the *old* payment reminder letter, the action-relevant information was hidden and spread out over a long, technical letter in the *old* design, also containing information that was only relevant for internal use. The quality of the *old* filing reminder letter was arguably in between. In the payment reminder experiment, the simplified presentation increased tax compliance by as much as 23% before the start of follow-up enforcement. This effect is larger than in the filing reminder experiment (8%) and an order of magnitude larger than in the payment experiment (0.7%). Hence, simplification was effective everywhere, but had a larger impact in contexts where the *old* letter was more complex.

The dynamic patterns discussed before, with larger effects at the deadline and after receipt of the letter, suggest that the simplified communication is effective in making the deadline more salient and reduces chances to forget to pay or file before. This is confirmed when considering the effects on the extensive and intensive margin. In particular, Panel A of Appendix Table C7 shows the treatment effects in the tax payment experiments (TP and TPR) on the fraction of the tax liability paid conditional on paying. We find positive effects of simplification at the intensive margin, but of much smaller magnitude than the extensive margin effects (and only significant in TP). The tax payment experiments (TP and TPR) also included treatments varying the personalization of the letter design. Specif-

ically, in the TP experiment, some simplified letters did not address taxpayers by name (Simplified Not Personalized), and in the TPR FY2015 experiment, in some letters with a deterrence message the female partner in a couple was addressed before the male (Explicit Penalty FM). These variations did not make any difference (see Appendix Table C6).

Deterrence While prior work - both theoretical and empirical - has highlighted the importance of deterrence to tackle tax evasion, our experiments show that making penalties explicit in tax correspondence can improve timely tax filing and payment too, with compliance effects between 0.5 and 3pp across the different experiments. We briefly discuss here the specific deterrence treatments and refer the reader to Appendix Table C1 for the exact wording of the messages. The baseline deterrence treatment in the tax payment and payment reminder experiments states the average penalty (of €209) explicitly. In the filing reminder experiment, the treatment effect is somewhat larger when instead of the average penalty the deterrence message states the range of possible penalties (from €5 to €1,250) and tax rate increases (from 10 to 200%). We also find that making enforcement explicit by emphasizing the seizing of income/assets to actually collect penalties further increased compliance.[20] We also tested a more implicit variation of the enforcement message, which emphasized that not paying taxes would be seen as an active choice, building on Hallsworth et al. (2015). This treatment had no significant effect, potentially in line with the ineffectiveness of the tax morale treatments in our context. In contrast, a message that empasized that by taking immediate action, taxpayers could avoid penalties significantly increased compliance. In the payment reminder experiment, making the penalty explicit in combination with the immediacy message increased compliance from 1pp to 1.7pp (see TPR, FY2015 in Appendix Table C6).[21] Also in the tax payment experiment, we ran a treatment in which we highlighted the returns to immediate action to avoid enforcement measures, which increased the treatment effect from the simplified letter from 0.4 to 0.7pp (see TP in Appendix Table C6). This complements the earlier finding from the simplification treatment that besides making the relevant information salient, there is also a role for encouraging immediate action. We do not find an effect of deterrence at the intensive margin, when looking at the paid tax liability conditional on paying (Appendix Table C7).

[20]The Explicit Penalty+Enforcement message increases compliance 2.5pp against 1pp for the Explicity Penalty message in TPR, FY2015 - see Appendix Table C6. The difference between the two coefficients is significant with a p-value of 0.001.

[21]The difference in treatment effects between the explicit penalty and the explicit penalty+immediacy treatment is significant with a p-value of 0.077.

Tax Morale Our finding that tax morale messages are ineffective in raising tax compliance contrasts with some earlier studies on tax payment (e.g., Hallsworth et al. (2017) in the UK) and on tax filing (e.g., Bott et al. (2017) on foreign income reporting in Norway). However, a series of studies have found no effects when introducing normative appeals (e.g., Blumenthal et al. (2001), John and Blume (2018)). We both widen and strengthen the evidence by finding no or negative results at the payment and the filing stage, for the full population of tax payers / filers and on the subset of late filers / payers. Since we work on the universe of Belgium tax payers, the estimates are sufficiently precise to reject at usual significance levels that tax morale messages have effects of a magnitude comparable to the simplification and deterrence treatments. The tax morale message is also consistent across different treatment variations used in previous papers, either emphasizing the social value of the tax expenditures, or invoking the social norm of tax compliance by other Belgian taxpayers. For the online tax filing experiment, the treatment is somewhat different (i.e., the pop-up of a pie chart of tax expenditures) and so is the compliance measure (i.e., reported taxable income). However, the conclusions are the same.[22]

Tax morale messages may be ineffective because the messages were ineffective at raising tax morale, or because tax morale itself is not an important driver of tax compliance. To shed some light on the reasons why tax morale messages are ineffective, we draw from the large-scale survey implemented in combination with the online TF experiment. Taxpayers were invited to participate to an online survey immediately after they filed. The response rates were similar in treatment and control (resp. 5.15% and 5.14%): in total 79,334 tax filers completed the survey. Appendix Table C8 presents treatment effects on survey responses. As expected, tax filers who had seen the pie chart were more likely to say that they knew how taxes were spent (column 1) and were indeed closer to the truth when asked about the share of government spending in each category (column 2).[23] Second, treated taxpayers did not only know better, they also agreed more with how taxes were spent in general (column 3). When asked to rank expenditures categories in terms of which the government should give priority to, their stated preferences were closer to the actual ranking (column 4). They also reported attaching more value to public services financed with tax revenues (column 5). In the end, however, treated tax filers were not more likely to be satisfied with

[22]Panel B of Appendix Table C7 shows the impact of the pie chart treatment on five other tax compliance outcomes, including self-employed profits and deductible expenses. The average treatment effect on tax compliance is precisely estimated, but always insignificant.

[23]Using respondents' responses, we construct a knowledge index equal to minus the standardized sum of absolute deviations between the stated and the actual share over all spending categories.

the general tax system and not more likely to agree with the statement that taxes should be reported honestly (column 6 and 7). These results suggest that while the pie chart treatment was effective in improving taxpayers' knowledge and appreciation of how their taxes were spent, it fell short of improving their tax morale.

Heterogeneous Effects Average treatment effects can mask important heterogeneity, which is important to better target interventions, and to gauge the distributional consequences of interventions that alleviate heterogeneous frictions.[24] We focus on the payment reminder experiments, for which we were able to obtain a large set of observables (including various demographics like age, family composition, region, amount owed, taxable income and solvency score). To discipline our analysis of treatment effect heterogeneity, we use the causal forests algorithm created by Wager and Athey (2018).[25]

Figure 3.4 plots the dispersion of the treatment effects by treatment category (bin size is set to 0.5pp for all figures). While the figure only uncovers the heterogeneity in treatment effects based on observables, it is interesting to compare the predicted heterogeneity across treatments using the same set of observables. Indeed, we see a wide dispersion for the simplification treatment, but less so for the deterrence and tax morale ones. Moreover, the effect of the simplification treatment never turns negative, while the deterrence treatment has negative effects for some tax payers. Interestingly, the tax morale treatments seem to backfire for most taxpayers: almost all estimated treatment effects are negative.

Using the same causal forests estimates, we can determine which observable characteristics drive the heterogeneity in treatment effects. Figures C3a to C3h in Appendix present the average of the different observables in each treatment effect quintile. The machine learning results identify four relevant dimensions of treatment heterogeneity: age, number of children, tax liability and solvency. We confirm that these dimensions matter everything else equal, by regressing tax compliance on interactions of the treatment with these four main characteristics, including interactions of the treatment with all other characteristics as controls. Table C9 presents the results.[26] Simplification is more effective among taxpayers with children, who may have a harder time to track deadlines. Simplification is

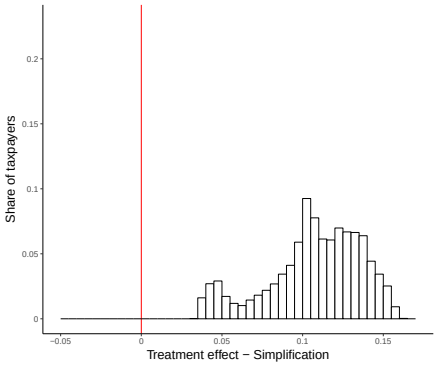
[24] See for example Alcott et al. (2018) in the context of using corrective sin taxes.

[25] According to Chernozhukov et al. (2018), we are in the case where the Wager and Athey (2018) method provides robust results: we have 10 dimensions of heterogeneity and about 230,000 observations ($\log(230,000) = 12 > 10$).

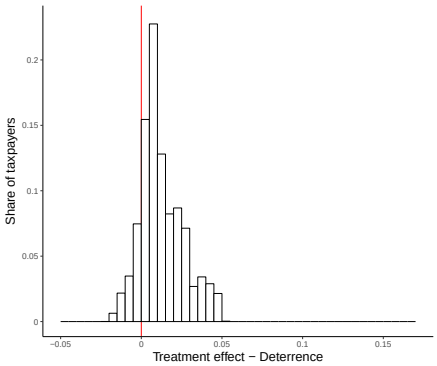
[26] Appendix Table C10 present similar results for the second TPR experiment (fiscal year 2015).

Figure 3.4: Distribution of treatment effects

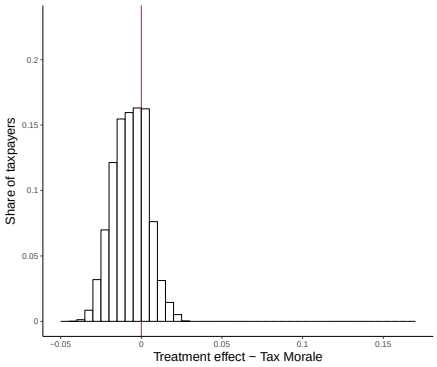
(a) Simplification



(b) Deterrence



(c) Tax Morale



Note: The figure presents the distribution of estimated treatment effects in the TPR FY2014 experiment. It uses the generalized random forest (GRF) algorithm (Wager and Athey, 2018) as described in the text. Figures (a)-(c) differ in the definition of treatment and control groups. In Figure (a) the control is composed of taxpayers who received the old letter and the treatment of taxpayers who received a simplified letter without any additional message. In Figure (b) and (c) taxpayers who received a simplified letter without any additional message are the control group. In Figure (b) the treatment is composed of taxpayers who received a simplified letter with a deterrence message. In Figure (c) the treatment is composed of taxpayers who received a simplified letter with an added tax morale message.

also more effective among taxpayers with a solvency score (as predicted by the tax administration) that is neither too high nor too low, i.e. it has little effect on people who pay their taxes readily or on people who face financial difficulties in

paying their taxes. Deterrence is most effective for younger taxpayers (who may be less aware of enforcement actions) and taxpayers with a lower outstanding liability (for whom the average penalty may seem high as compared to what they owe). There is no obvious pattern for gender, language, region or income.

3.5 Simplification and Enforcement

The previous section compared the effect of different letter interventions on tax compliance. As shown in Section 3.2, we eventually care about how much the interventions increase tax revenues and reduce the need for follow-up enforcement by the tax authority. This section estimates the cost-effectiveness of letter interventions relative to standard enforcement actions. To that purpose, we exploit a regression discontinuity in enforcement intensity for the late tax payers, which, combined with the experimental design of the tax payment reminders, provides a unique opportunity to compare the compliance effect of letter interventions and standard policy levers for the same population and in the same setting.

3.5.1 Nudges vs. Enforcement

The tax administration relies on various enforcement actions to make late payers comply. The first follow-up intervention for late tax filers and taxpayers is naturally the reminder letter, which we experimentally manipulated. Individuals who do not comply after receiving the reminder are subject to further enforcement actions. Local tax administrators have some discretion in the choice of enforcement mechanisms. Commonly used tools for payment non-compliers include sending registered letters (which require confirmation of receipt), imposing garnishments and the use of bailiffs. The dynamic pattern of the treatment effects (Figure 3.3) showed that the letter treatments accelerated tax payments, but that their final effect on tax compliance was more modest. The timing of the decline in treatment effects corresponds to the start of the enforcement actions undertaken by the administration, which suggests that these actions are responsible for the control group catching up with treatment.

To provide causal evidence on the effect of enforcement actions, we implement a regression discontinuity design which exploits exogenous variation in enforcement intensity at a specific threshold for the outstanding tax liability. We then combine the regression discontinuity with the simplification treatment to understand both how much the simplification treatment reduced the need for follow-up enforcement and how much the follow-up enforcement reduced the impact of the

simplification treatment.

As Panel (a) of Figure 3.5 shows, there is a clear jump in the probability of enforcement actions above the tax liability threshold (normalized to 0 for confidentiality reasons), both in the treatment and control group.[27] There is no evidence of bunching below the threshold, which confirms that it is not known to the public (see Figure C4). Moreover, before enforcement started, the probability of paying is smooth at the cut-off in both groups. This probability of paying, however, is much higher in the treatment than in the control group, which explains why both to the left and to the right of the cut-off, the treatment group is less likely to be subject to enforcement interventions. Importantly, the absence of discontinuities in the density and the pre-enforcement outcomes, both in the treatment and control group, seems to validate the use of a regression discontinuity design to estimate the causal effect of enforcement actions.

The impact of enforcement on compliance is illustrated in panel (b) of Figure 3.5. The fraction of taxpayers who have paid after 180 days is higher to the right than to the left of the threshold. Interestingly, compliance levels are similar in the treatment and control group to the right of the cut-off where enforcement intensity is high, while to the left where intensity is lower the treatment group is substantially more compliant.

To estimate the causal effects of the simplification treatments and the enforcement actions, we implement the standard regression discontinuity method in the control group, and add treatment dummies. Formally, let Y_i denote the tax compliance outcome of individual i, z_i their tax liability, c the tax liability cutoff. As before, S_i a dummy variable equal to one for the randomly assigned group who received the simplified letter and X_i is a vector of individual characteristics (see Table 3.1). The estimating equation is:

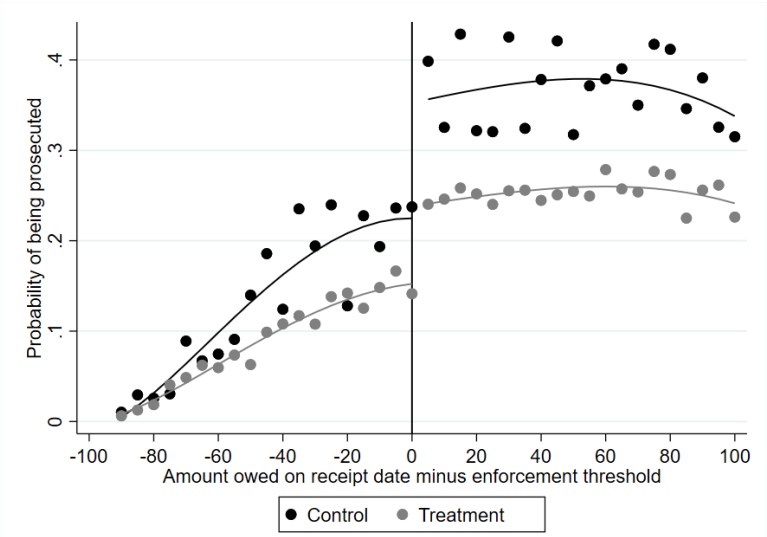
$$\begin{aligned} Y_i &= \alpha + \beta_S S_i + \beta_E 1\{z_i - c > 0\} + \beta_{S,E} S_i \times 1\{z_i - c > 0\} \\ &+ \delta_{C,l}(z_i - c) + \delta_{C,r} 1\{z_i - c > 0\} \times (z_i - c) + \delta_{S,l} S_i \times (z_i - c) \\ &+ \delta_{S,r} S_i \times 1\{z_i - c > 0\} \times (z_i - c) + \gamma X_i + \varepsilon_i \end{aligned}$$

Due to the random assignment, β_S identifies the effect of simplification at the cut-off from the left, where enforcement is weaker. Due to the regression-discontinuity, β_E identifies the effect of additional enforcement actions on tax compliance in the control group. Combining the two sources of variation, $\beta_{S,E}$ identifies the differ-

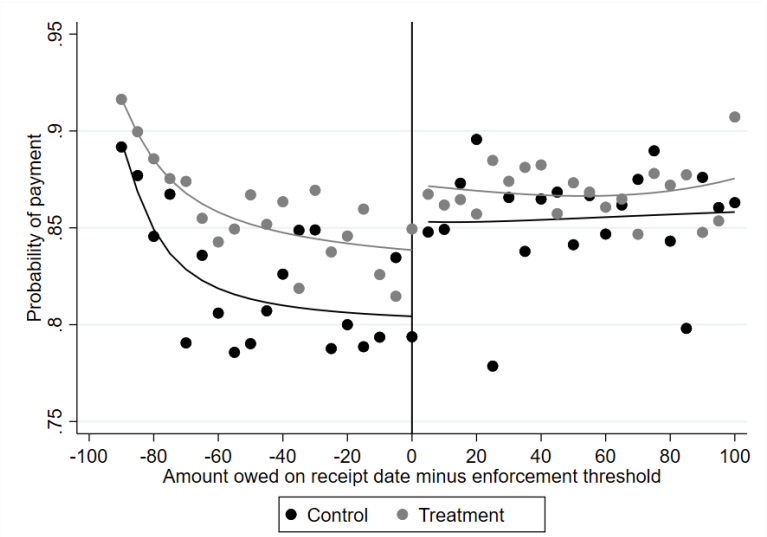
[27] We exclude taxpayers with a liability exactly at the cut-off. The threshold value is a round number and the distribution of liabilities shows bunching at all round numbers in the vicinity of the threshold.

Figure 3.5: Effects of Enforcement and Simplification

(a) Probability of Enforcement after 180 Days



(b) Probability of Partial Payment after 180 Days



Note: The figure is based on the TPR FY2014 experiment. It shows probability of enforcement after 180 days (Panel (a)) and probability of partial payment after 180 days (Panel (b)) by initial amount owed (centred at the enforcement threshold). Bin size is set to €5 and amounts within €100 of the enforcement threshold are considered. Fractional polynomial predictions plotted.

ence in treatment effects due to higher enforcement at the threshold. As in a typical regression discontinuity setting, $\delta_{C,l}$ and $\delta_{C,r}$ capture the relation between the forcing variable (tax liability) and the outcome (tax compliance) to the left and the right of the discontinuity, while $\delta_{S,l}$ and $\delta_{S,r}$ allow this relation to be different for the treatment group. An alternative interpretation is that the latter interaction terms allow for heterogeneity in treatment effects depending on the tax liability, both to the left and to the right of the cutoff.

Table 3.4 presents the corresponding regression results, using the Imbens-Kalyanaraman bandwidth computed for the control group in our experiment. We first consider the RDD estimates for the control group in our experiment. Column 1 confirms that the probability of enforcement increased by 15pp, from 21 to 36%, at the threshold. Before enforcement actions begun, the payment probability, however, was smooth at the threshold (Column 2). In contrast, 180 days after reminder receipt, the payment probability increased by 6.1pp at the threshold, reaching a probability of 87% for taxpayers in the control group to the right of the threshold (Column 3). Second, we consider the effects of simplification, not just on payment, but also on follow-up enforcement. As Column 1 shows, simplification decreased the probability of any enforcement action by almost half, from 21% in the control to 13%. This is due to the fact that simplified reminders made late payers 15pp more likely to pay before enforcement actions begun: from 49 to 64% (Column 2). Note that these effects are larger than those we report for the whole late payer sample (see Table 3.2). After 180 days, once payment rates in the control group have increased to 81%, the treatment effects were smaller, but still significant: a 4.4pp increase (Column 3). Finally, we estimate the difference in treatment effects to the left and to the right of the threshold. While the difference $\beta_{S,E}$ is not significant, the estimate is negative and large enough to mostly offset the positive treatment effect on the probability of paying at 180 days (Column 3).[28] This confirms the graphical evidence that with high intensity enforcement the effects of simplification in the long run are virtually zero.

Table 3.4: RDD: Effect of Simplification vs. Enforcement in TPR 2014

	Probability of enforcement at 180 days	Probability of some payment at 14 days (before enforcement)	Probability of some payment at 180 days (after enforcement)
	(1)	(2)	(3)
Simplified (S)	-0.078	0.151	0.044
	(0.025)	(0.025)	(0.019)
Enforcement	0.146	0.006	0.061
	(0.034)	(0.034)	(0.027)
S * Enforcement	-0.064	0.000	-0.027
	(0.036)	(0.036)	(0.028)
Control Mean	0.210	0.489	0.813
N	16,277	23,312	21,894

Note: The table presents simplification treatment effect estimates and enforcement RDD estimates for the TPR experiment (FY2014). Simplified is a dummy variable equal to one for taxpayers who received a simplified letter. Enforcement is a dummy variable equal to one for liability amounts above the cut-off value. Control variables are listed in Table 3.1. Standard errors in parentheses are clustered by date of letter receipt.

While the compliance benefits of nudges seem to disappear because of follow-up interventions on non-compliant taxpayers, they do bring important benefits

[28] Note that these effects are driven by registered letters and garnishments (Appendix Table C12).

by saving on enforcement costs as we discuss further below. Interestingly, we can also use our results to estimate the counterfactual effect of simplification after 180 days if the follow-up enforcement intervention had not taken place. Of course, in practice, the reminder letters effectiveness depends on tax payers' expectation of the follow-up enforcement by the administration. Still, to calculate the effect of simplification net of the crowd-out by the follow-up interventions, we impute the level of compliance based on the difference in compliance between high and low intensity enforcement groups scaled up by the difference in enforcement probability between them. Formally, let Y denote the payment probability, F the enforcement probability, z tax liability, c the cutoff and S letter simplification. Let the superscript F and Y denote the estimated coefficients when the dependent variable is F and Y, respectively. We approximate the average treatment effect in absence of enforcement, ATE_0, by:

$$\begin{aligned} ATE_0 &\approx \left[E(Y|_{S=1,z<c}) - E(F|_{S=1,z<c})\frac{E(Y|_{S=1,z>c}) - E(Y|_{S=1,z<c})}{E(F|_{S=1,z>c}) - E(F|_{S=1,z<c})}\right] \\ &- \left[E(Y|_{S=0,z<c}) - E(F|_{S=0,z<c})\frac{E(Y|_{S=0,z>c}) - E(Y|_{S=0,z<c})}{E(F|_{S=0,z>c}) - E(F|_{S=0,z<c})}\right] \\ &= \left[\left(\widehat{\alpha^Y} + \widehat{\beta_S^Y}\right) - \left(\widehat{\alpha^F} + \widehat{\beta_S^F}\right)\frac{\left(\widehat{\beta_E^Y} + \widehat{\beta_{S,E}^Y}\right)}{\left(\widehat{\beta_E^F} + \widehat{\beta_{S,E}^F}\right)}\right] - \left[\widehat{\alpha^Y} - \widehat{\alpha^F}\frac{\widehat{\beta_E^Y}}{\widehat{\beta_E^F}}\right] = 0.077 \end{aligned}$$

This calculation relies on a homogeneity assumption: we need that the effect of enforcement on the payment probability is the same for taxpayers who pay only when enforcement intensity increases from below to above the threshold and for taxpayers who pay even with low intensity enforcement. The counterfactual analysis suggests that in absence of the follow-up enforcement actions, the effect of simplification on the payment probability of late payers would have been 7.7pp after 180 days, which is approximately half of the effect estimated before enforcement actions begun (15pp).

3.5.2 Cost-Effectiveness and Welfare

We now evaluate the cost-effectiveness of the simplification treatment. We consider three closely related approaches. First, we compare the benefits of the treatment in terms of additional revenue and savings on enforcement actions to the costs of simplifying the tax correspondence. Second, we compare the cost of raising one euro of extra revenue through reminder simplification and through enforcement actions. Finally, we calculate the total cost of enforcement actions that is needed to raise the same extra revenue as the simplification treatment

could.

The first method is based on experimental results only. To compute extra revenues, we estimate the effect of simplified letters on the probability of paying taxes as late as possible in the tax cycle, which is 180 days after the payment deadline, and assume that after this date the treatment effect will remain constant.[29] As Table 3.3 shows, the estimated treatment effect on the probability of payment at 180 days is 1pp, which we multiply by the average amount paid, conditional on a payment, at that date (€1,615) and the number of tax payers in the treatment group (205,014) to obtain total extra revenues equal to €3.16 million. To compute savings on the cost of enforcement, we estimate the effect of simplified letters on the probability of the three most common forms of enforcement actions – registered letters, garnishment and bailiffs. Multiplied by the cost of the respective enforcement measures, we obtain a total cost saving of €0.70 million.[30] Adding the extra revenues and costs savings on enforcement, the total benefit of the intervention equals €3.86 million. In comparison, the costs of simplification were negligible: the administration paid €69,300 for the design of the new letter, including ICT staff, data analysts, legal experts, communication staff and management, and the printing of the new (colored) letter costs an extra €0.05 per letter. The total cost of simplifying the reminder letters amounts to €79, 550 and is about 50 times smaller than its benefits. Simplifying the reminder letters was thus a high return investment for the tax administration.

The second method builds on the regression discontinuity results from the previous section. Since we are able to estimate the compliance effects of the simplification treatment and the enforcement interventions separately, we can ask what the most cost-effective way is to raise one euro of extra revenue. The conceptual framework in Section 3.2 made clear that from an efficiency prespective, an optimal use of simplification and enforcement actions by the government should equalize the marginal cost of raising an additional euro of revenue between them. For the enforcement interventions, we first use regression discontinuity estimates for the increase in the probability that registered letters (11.0pp) and garnishment (7.1pp) were sent at the threshold (see Appendix Table C12) and their cost (€5.7 and €17.1 respectively) to compute the cost of the increase in enforcement

[29] After 180 days, tax filing for the next fiscal year begins: the administrative data that we use does not allow us to track outstanding debts separately from new tax liabilities.

[30] As Appendix Table C11 shows, the estimated treatment effects on follow-up enforcement are −7.4pp for registered letters, −2.8pp for garnishment actions and −1.2pp for bailiffs. Multiplying these figures by the cost of each action and the number of treated taxpayers, we obtain costs savings of €86, 436 for registered letters, €97, 357 for garnishment and €517, 318 for bailiffs.

intensity at the threshold, which is €1.85.[31] We then use regression discontinuity estimates of the effect of enforcement intensity on the probability of payment at 180 days (from Table 3.4) multiplied by average payments made at the threshold to estimate additional revenues raised. The ratio of the two, i.e., the cost of raising one more euro of tax revenues through enforcement is equal to €0.31. This estimate is arguably in the range of standard estimates of the marginal excess burden of personal income taxes, suggesting that the enforcement intensity may well be desirable (Keen and Slemrod, 2017). In comparison, the resource cost of using nudge interventions is much smaller: €79,550 in total, or €0.39 per letter sent. We multiply our counterfactual estimate of the effect of simplification on the probability of payment in the absence of follow-up enforcement by the average tax payment, and obtain €7.53 extra revenue per letter. Hence the cost of raising one euro with simplified reminders is €0.05, which is six times smaller than with enforcement actions.[32] This second method confirms that simplifying reminders is far more cost-effective than intensifying enforcement.

The third method extrapolates the regression discontinuity results to the whole sample, using a back-of-the envelope calculation. At the enforcement threshold, the treatment effect was 15.1pp after 14 days and the counterfactual effect absent follow-up enforcement at 180 days was 7.7pp (Table 3.4). Hence for the whole sample the estimated treatment effect of 10.3pp after 14 days suggests that the counterfactual effect, in the absence of follow-up enforcement, would have been $10.3 * 7.7/15.1 = 5.2$pp at 180 days. Multiplying this figure by the amount paid by the average taxpayer and by the number of letters sent gives €17.5million of extra revenue. To obtain these extra revenues with traditional enforcement methods at the cost of 31 cents per euro raised, the government would have had to spend €5.4 million. This is again subtantially higher than the cost of the simplification intervention (€79,550).

Regardless of the method we use for the cost-benefit analysis, simplifying letters seems highly cost effective, in itself and when compared to the alternative of using standard enforcement actions. The above calculations, however, ignore other welfare-relevant considerations that may be important when assessing the use of nudges. First of all, the letter treatments - when successful - changed the net transfers between taxpayers and the government, not only by affecting the

[31] As Appendix Table C12 shows, there is no significant increase in the use of bailiff at the threshold. As an enforcement tool, the use of bailiffs is applied to debts of relatively large amounts, while registered letters and garnishments are more often employed.

[32] We consider this a conservative estimate as the cost of nudging is largely driven by the fixed costs of experimental design. If these are ignored the per letter cost goes down to 0.05 making it eight times cheaper and thus lowering significantly the cost to benefit ratio of the nudging intervention.

taxes paid, but also avoiding the late penalties and interests on outstanding tax liability. Second, the nudges can affect individuals' welfare above and beyond their after-tax income. The simplified correspondence reduces compliance costs, but may also reduce the disutility of paying taxes.[33] While the same may be true for highlighting the public value of taxes paid, the opposite effect seems as plausible when using deterrence or invoking social norms.

3.5.3 Long-term Effects

We have shown that simplification is effective at different stages of the tax process, and for different subpopulations of income taxpayers. We have also shown that in the case of payment reminders, it is very cost effective, in itself and as compared to traditional enforcement actions. We now ask whether the simplification intervention only works once and its effects are short-lived, or to contrary, (i) has long-term effects and (ii) can be used repeatedly on the same taxpayers. To test this, we exploit the two payment reminder experiments carried out over two consecutive years.

We first investigate whether simplification of communication in one fiscal year can improve compliance in subsequent years. We use the randomization in the FY2014 payment reminder experiment to estimate the effect of reminder letters on timely payment in the next two fiscal years (FY2015 and FY2016). The results are shown in column 1 of Panel A of Table 3.5. We find a positive and significant effect of simplification on tax compliance in the next financial year. The probability of paying taxes on time in FY2015 increased by 1.3pp. Note that this long-term effect of simplification of the reminder letter is twice as large as the short-run effect of the simplification of the tax bill itself (0.5pp increase in the probability of meeting the deadline, see column 1 in Table 3.2). This may be due to the fact that the simplification of the reminder letter was more substantial than the simplification of the tax bill, as discussed in the previous section. Also, the reminder letters were sent to a subsample of taxpayers who may be more sensitive to simplification. Two fiscal years after the intervention, the effect of simplification had declined to 0.5pp, and the coefficient is no longer significant (column 2 Panel A of Table 3.5). In contrast with simplification, the deterrence messages had no effect in the following fiscal years, but the negative effect of tax morale messages was remarkably persistent. Overall, these results suggest that small nudges can have long-term effects, and that the benefits of simplification

[33]For example, Di Tella et al. (2015) show that complexity can lead people to be "conveniently upset" and use it as an excuse not to comply.

may be even larger than our cost-benefit analysis based on the effects in one fiscal year only would suggest.

Table 3.5: Long-term and Repeated Treatment Effects

Panel A: Long-term Effects	Probability of being on time with payment FY+1 year (1)	Probability of being on time with payment FY+2 years (2)
Simplified (S)	0.013	0.005
	(0.003)	(0.004)
+ Deterrence	-0.002	-0.003
	(0.003)	(0.002)
+ Tax Morale	-0.009	-0.005
	(0.002)	(0.003)
P-values of tests:		
Simplified=Control	0.001	0.430
S+Deterrence=Simplified	0.493	0.509
S+Tax Morale=Simplified	0.016	0.253
Control mean	0.703	0.776
N	229,751	229,751
Panel B: Repeated Treatment	Probability of some payment at 14 days (before enforcement) in FY2015	
	Sample of Taxpayers late in FY2014 and FY2015 (1)	Sample of Taxpayers late in FY2014 (2)
Simplified 2014	-0.000	-0.001
	(0.010)	(0.005)
Simplified 2015	0.099	0.024
	(0.011)	(0.007)
S 2014 * S 2015	-0.002	0.004
	(0.009)	(0.006)
P-values of tests:		
Simplified 2014=Control	0.424	0.956
Simplified 2015=Control	0.001	0.001
S 2014*S 2015=S 2015	0.535	0.278
Control mean	0.410	0.825
N	66,705	229,751

Note: The table presents results from the replication, long-term and repeated treatment analysis. The sample in Panel A is the universe of late payers in FY2014. In Panel B column 1 it is composed of taxpayers who were late with payment in both FY2014 and FY2015. In Panel B column 2 it is composed of the universe of late payers in FY2014. Control variables are listed in Table 3.1. Standard errors in parentheses are clustered by date of letter receipt. p-values are adjusted for multiple hypothesis testing (List et al., 2016).

We then ask whether repeated interventions remain effective. For this we use the cross-randomization of the FY2014 and FY2015 payment reminders experiments. First, we check that the FY2014 experimental results replicate in FY2015 (see Appendix Table C13). In FY2015 as in FY2014, simplifying tax reminders had a large positive effect on the probability of paying before enforcement starts (+10.7pp), and deterrence messages had an additional positive effect (+1.4pp), while tax morale messages had a negative effect (−1.2pp). Interestingly, mixing deterrence and tax morale messages had a significantly smaller impact than deterrence messages alone. Given that the treatment effects replicate, we can

now test whether simplified letters had a larger or smaller effect for taxpayers who received them twice, i.e. whether repetition induced a reinforcement or a fatigue effect. The results are presented in Panel B of Table 3.5. To simplify the exposition, we estimate the effect of receiving *any* simplified letter in FY2014, in FY2015 or in both years.[34] Among taxpayers who were late twice, the estimated effect of the simplified letter in FY2015 is again large (9.9pp) and comparable to the estimated effect of a simplified letter on the late tax payers in FY2014 (Table 3.2) and in FY2015 (Appendix Table C13). Interestingly, among taxpayers who were late twice, but already received a simplified letter in FY2014, the effect of the simplified letter in FY2015 is the same (i.e., the interaction coefficient is zero). As simplification is not less effective when used repeatedly, this result suggest that fatigue effects are unimportant. However, taxpayers who were late twice are of course a selected sample of taxpayers and we know that the simplification in FY2014 itself affected the selection as it decreases the probability of being late again in FY2015. For completeness, we also report effects for the whole sample of taxpayers who were late in FY2014 (column 2) rather than just for the subsample (30%) of taxpayers who were late again in FY2015 (column 1). Also in the whole sample, we find a significant positive effect of simplification in FY2015 (2.4pp) and no evidence that simplification was less effective for those who had received the simplified letter in FY2014 (precise zero on the interaction term).

3.6 Conclusion

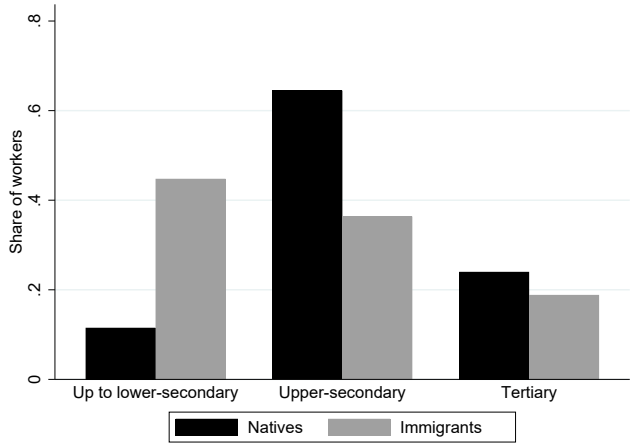
Based on a series of population-wide experiments in Belgium, we show that simplifying communication by the tax administration consistently improves tax compliance. Simplification makes taxpayers pay taxes on time and makes both late filers and late payers comply more swiftly than they would otherwise. Our results also demonstrate the added benefits from including deterrence messages in the same context but suggest that invoking tax morale does not raise compliance and often backfires. Finally, we estimate causally the costs and benefits of simplification as compared to traditional enforcement actions, and find simplification to be highly cost effective. The positive effects of simplification persist in the next fiscal year and are sustained when simplification is repeated. Making it as easy as possible to comply therefore deserves greater attention from tax administrations around the world.

[34]The estimated effects of simplification are similar when we include dummy variables for the different messages and their interactions (Appendix Table C14). Note that with the treatment interactions, the estimation is based on relatively small subsamples and the estimates become less precise.

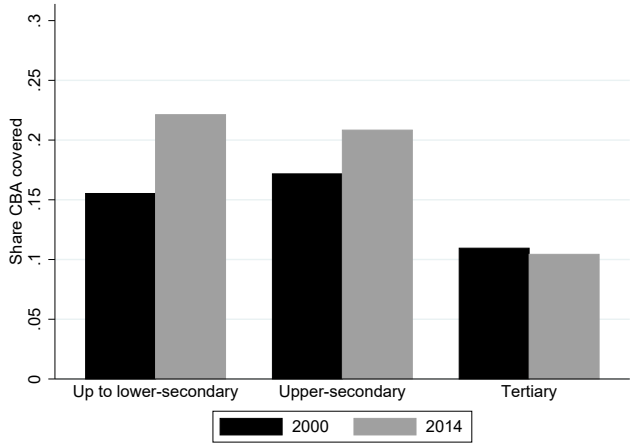
Appendix A (for Chapter 1)

Figure A1: Educational attainment and CBA coverage

(a) Educational attainment of native and migrant workers

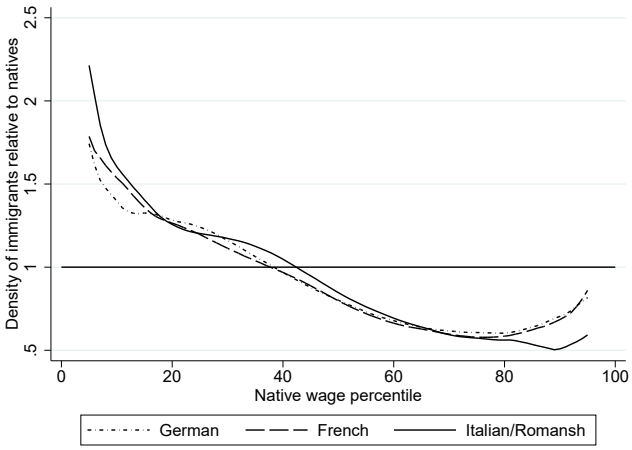


(b) Native CBA coverage by educational attainment



Note: Figure (a) presents the share of native and immigrant workers by highest level of educational attainment for the period 2000–2014. Lower-secondary level of education is compulsory education as highest degree, upper-secondary is an apprenticeship or a matura, tertiary is a degree from a university, university of applied sciences, university of teacher education or a professional degree. Figure (b) presents the share of native workers employed in an industry with a CBA by highest level of educational attainment. Source: SESS.

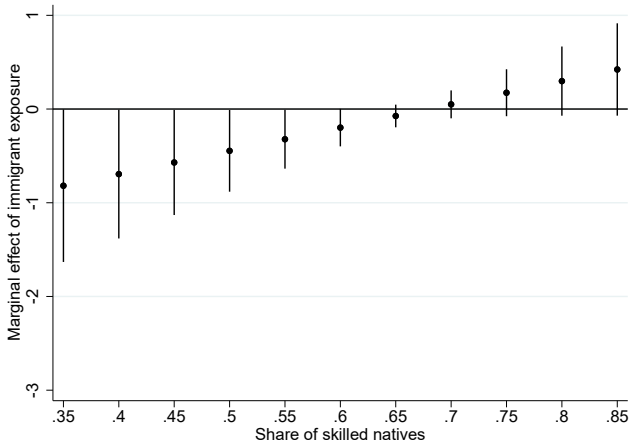
Figure A2: Position of foreign workers in native wage distribution by language region



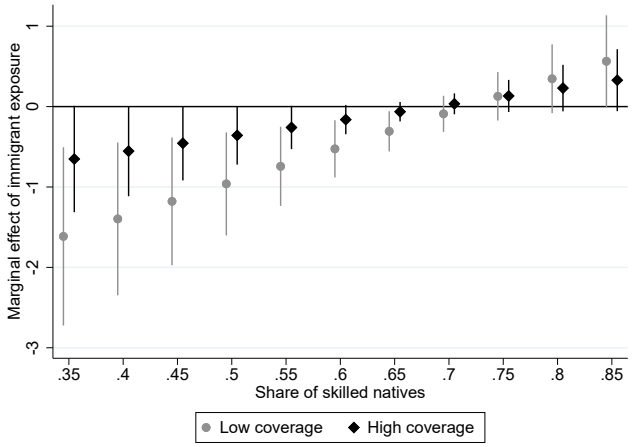
Note: The figure present kernel estimates of the density of foreign workers along the native wage distribution by language region. Data is pooled from 2000 to 2014. Source: SESS.

Figure A3: Voting analysis by continuous native educational attainment

(a) Overall estimates



(b) Estimates by level of CBA coverage



Note: The figure presents estimates from an IV regression using municipality level data. The outcome is the share of pro-immigrant votes. Share of migrants is the number of foreign residents in municipality and year divided by native population in municipality in 2000. Share of skilled natives is the share of native residents with upper-secondary or higher level of education in 2000. Controls are listed in Table A2; all specifications include municipality and vote fixed effects. Weights assigned to observations reflect the number of Swiss residents in 2000. Standard errors are clustered at the municipality level, 95% confidence intervals plotted. In Figure (b) effects at the 10th and 90th percentile of the coverage measure are reported. Source: FSO, SECO, ZEMIS.

Table A1: Votes on immigration policy 2000–2014

Nr.	Date	Title	Type	Category	Impact	Approval (%)
464	21.05.2000	Bilateral agreements with the EU	Optional	European politics	Positive	67.2
467	24.09.2000	Initiative "for regulation of immigration"	Initiative	Immigration policy	Negative	36.2
519	25.09.2005	Extension of the agreement for free movement of persons (AFMP)	Optional	European politics	Positive	56.0
524	24.09.2006	Federal Act on Foreign Nationals (AuG)	Optional	Immigration policy	Negative	68.0
540	08.02.2009	Approving the continuation of the AFMP and extension to Bulgaria and Romania	Optional	European politics	Positive	59.6
580	09.02.2014	Initiative "against mass immigration"	Initiative	European politics	Negative	50.3

Note: Table presents the list of votes considered in the analysis. Types of votes are compulsory referendum, optional referendum, popular initiative and counter proposal. Impact refers to the expected effect of the vote on the level of immigration in Switzerland. We classify proposals with an expected positive impact as "pro-immigration" and proposals with an expected negative impact as "anti-immigration" votes.

Table A2: Summary statistics of control variables

	N	Mean	Sd	Min	Max
Swiss Earnings Structure Survey (SESS)					
Share women	848	0.376	0.053	0.070	0.690
Mean age	848	40.623	1.059	31.394	46.135
Share above lower-secondary educated	848	0.875	0.043	0.429	0.985
Swiss Labor Force Survey (SLFS)					
Share women	1590	0.509	0.041	0.169	0.839
Mean age	1590	42.113	1.318	33.185	51.463
Share above lower-secondary educated	1590	0.866	0.042	0.507	1.000
Federal Statistical Office					
Share women	8860	0.504	0.016	0.331	0.606
Mean age (2000)	2215	39.099	2.393	29.487	58.655
Share above lower-secondary educated (2000)	2215	0.713	0.069	0.313	0.872

Note: Table presents summary statistics for control variables from the SESS and SLFS surveys, and the 2000 census. In the wage analysis weights equal the number of native workers in 2000 (SESS data), in the employment analysis the number of native respondents 18-65 years of age (SLFS), the number of Swiss residents in municipality in 2000 for census data. SESS and SLFS variables are measured at the commuting zone level, census variables at the municipality level. Source: FSO, SESS, SLFS, ZEMIS.

Table A3: Voting behavior and stated preferences

	Outcome: equal to one if respondent voted pro-immigration, zero otherwise					
	(1)	(2)	(3)	(4)	(5)	(6)
Prefers equal opportunities	0.349	0.325	0.342			
	(0.015)	(0.016)	(0.016)			
Prefers open Switzerland				0.431	0.398	0.419
				(0.020)	(0.020)	(0.020)
Age		0.001	-0.002		0.000	-0.002
		(0.003)	(0.003)		(0.003)	(0.003)
Age squared		0.000	0.000		-0.000	0.000
		(0.000)	(0.000)		(0.000)	(0.000)
Female		0.060	0.036		0.066	0.040
		(0.015)	(0.015)		(0.015)	(0.015)
Teacher education		-0.168			-0.174	
		(0.018)			(0.018)	
Upper-secondary		-0.074			-0.087	
		(0.021)			(0.022)	
Lower-secondary		-0.223			-0.227	
		(0.032)			(0.033)	
Hh inc. 7001-9000 CHF			0.001			-0.018
			(0.022)			(0.022)
Hh inc. 5001-7000 CHF			-0.065			-0.087
			(0.021)			(0.021)
Hh inc. 3001-5000 CHF			-0.056			-0.079
			(0.023)			(0.023)
Hh inc. <3000 CHF			-0.138			-0.122
			(0.032)			(0.032)
N	3525	3525	3525	3549	3549	3549

Note: Table presents estimates from OLS voting analysis using individual-level data. All specifications include place of residence and vote fixed effects. We code an individual with a response 1–3 as in favor of equal opportunities / open Switzerland, and 4–7 as not in favor. Base category for education is tertiary, base category for household income is above 9000 CHF per month. Robust standard errors in parentheses. Source: Vox Survey.

Table A4: Voting analysis by native educational level: IV checks

	Outcome: share of pro-immigration votes					
	Without controls			With controls		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 1970						
Sh. migrants	0.256	-1.841	-0.140	0.004	-2.499	-0.406
	(0.140)	(0.850)	(0.151)	(0.134)	(1.526)	(0.215)
Sh. migrants x Sh. skilled		2.667			3.425	
		(1.177)			(2.173)	
Sh. migrants x Q2 sh. skilled			0.049			0.181
			(0.105)			(0.183)
Sh. migrants x Q3 sh. skilled			0.092			0.235
			(0.097)			(0.208)
Sh. migrants x Q4 sh. skilled			0.325			0.445
			(0.112)			(0.241)
Sh. migrants x Q5 sh. skilled			0.348			0.433
			(0.147)			(0.329)
First stage F-stat	18.725	9.500		14.531	7.610	
N	13290	13290	13290	13290	13290	13290
Panel B: Interaction 4-digit industry						
Sh. migrants		-4.041	-0.454		-3.553	-0.416
		(1.172)	(0.143)		(1.104)	(0.127)
Sh. migr. x Sh. CBA cov.		9.817	1.267		8.345	1.036
		(2.419)	(0.397)		(1.865)	(0.418)
Sh. migr. x Sh. skilled		5.744			4.987	
		(1.725)			(1.652)	
Sh. migr. x Sh. CBA cov. x Sh. skilled		-13.577			-11.490	
		(3.566)			(2.789)	
Sh. migr. x Q2 sh. skilled			0.054			0.013
			(0.137)			(0.143)
Sh. migr. x Q3 sh. skilled			0.209			0.187
			(0.157)			(0.154)
Sh. migr. x Q4 sh. skilled			0.643			0.556
			(0.147)			(0.137)
Sh. migr. x Q5 sh. skilled			0.858			0.695
			(0.277)			(0.298)
Sh. migr. x Sh. CBA cov. x Q2 sh. skilled			-0.228			-0.056
			(0.401)			(0.422)
Sh. migr. x Sh. CBA cov. x Q3 sh. skilled			-0.588			-0.486
			(0.426)			(0.430)
Sh. migr. x Sh. CBA cov. x Q4 sh. skilled			-1.245			-1.060
			(0.392)			(0.416)
Sh. migr. x Sh. CBA cov. x Q5 sh. skilled			-2.017			-1.621
			(0.685)			(0.572)
N	.	13284	13284	.	13284	13284

Note: Table presents estimates from OLS and IV regressions using municipality-level data. Share of migrants is the number of foreign residents in a municipality in a year divided by native population in a municipality in 2000. Share skilled is the share of native residents with upper-secondary or higher level of education in 2000. Controls are listed in Table A2; all specifications include municipality and vote fixed effects. Weights assigned to observations equal the number of Swiss residents in 2000. Standard errors in parentheses are clustered at the municipality level. Source: FSO, ZEMIS.

Table A5: Voting analysis by native educational level and CBA coverage: IV leave-one-out

	Outcome: share of pro-immigration votes					
	2000 I	2000 II	2005	2006	2009	2014
	(1)	(2)	(3)	(4)	(5)	(6)
Sh. migrants	-0.259	-1.273	-0.913	0.134	-0.927	-1.411
	(0.181)	(0.219)	(0.183)	(0.161)	(0.189)	(0.353)
Sh. migr. x Sh. CBA cov.	0.903	3.960	2.508	-0.702	2.471	3.147
	(0.570)	(0.687)	(0.561)	(0.491)	(0.554)	(0.995)
Sh. migr. x Q2 sh. skilled	-0.176	0.528	0.069	-0.272	0.190	0.110
	(0.198)	(0.251)	(0.208)	(0.184)	(0.219)	(0.352)
Sh. migr. x Q3 sh. skilled	0.207	0.902	0.535	-0.005	0.690	0.255
	(0.213)	(0.254)	(0.240)	(0.177)	(0.220)	(0.598)
Sh. migr. x Q4 sh. skilled	0.443	1.393	0.923	0.232	1.151	0.835
	(0.196)	(0.227)	(0.195)	(0.181)	(0.201)	(0.368)
Sh. migr. x Q5 sh. skilled	0.449	1.684	0.970	0.199	1.336	0.787
	(0.340)	(0.357)	(0.394)	(0.303)	(0.326)	(0.778)
Sh. migr. x Sh. CBA cov. x Q2 sh. skilled	0.636	-1.784	-0.290	0.981	-0.480	-0.417
	(0.618)	(0.761)	(0.580)	(0.542)	(0.612)	(0.830)
Sh. migr. x Sh. CBA cov. x Q3 sh. skilled	-0.534	-2.787	-1.604	0.196	-1.745	-0.588
	(0.668)	(0.809)	(0.692)	(0.530)	(0.640)	(1.430)
Sh. migr. x Sh. CBA cov. x Q4 sh. skilled	-0.807	-3.644	-2.294	-0.077	-2.550	-1.759
	(0.675)	(0.748)	(0.615)	(0.552)	(0.603)	(0.913)
Sh. migr. x Sh. CBA cov. x Q5 sh. skilled	-1.346	-4.143	-2.480	-0.018	-3.132	-1.940
	(0.853)	(0.954)	(0.965)	(0.624)	(0.724)	(1.528)
N	11075	11075	11075	11075	11075	11075

Note: Table presents estimates from IV regressions using municipality-level data. Share of migrants is the number of foreign residents in a municipality in a year divided by native population in a municipality in 2000. Share skilled is the share of native residents with upper-secondary or higher level of education in 2000. Controls are listed in Table A2; all specifications include municipality and vote fixed effects. Weights assigned to observations equal the number of Swiss residents in 2000. Standard errors in parentheses are clustered at the municipality level. Source: FSO, SECO, ZEMIS.

Table A6: Voting analysis by native educational level and CBA coverage: IV MS-region

	Outcome: share of pro-immigration votes			
	(1)	(2)	(3)	(4)
Sh. migrants	-9.522	-2.365	-6.764	-2.006
	(4.158)	(0.908)	(2.720)	(0.835)
Sh. migr. x Sh. CBA cov.	29.000	7.646	22.798	6.759
	(13.746)	(2.890)	(11.324)	(2.768)
Sh. migr. x Sh. skilled	12.365		8.556	
	(5.725)		(3.796)	
Sh. migr. x Sh. CBA cov. x Sh. skilled	-36.275		-28.496	
	(19.306)		(16.426)	
Sh. migr. x Q2 sh. skilled		1.125		1.372
		(0.862)		(0.812)
Sh. migr. x Q3 sh. skilled		1.245		1.025
		(0.915)		(0.851)
Sh. migr. x Q4 sh. skilled		1.916		1.528
		(0.996)		(0.924)
Sh. migr. x Q5 sh. skilled		2.189		1.744
		(1.016)		(0.937)
Sh. migr. x Sh. CBA cov. x Q2 sh. skilled		-3.228		-4.499
		(2.669)		(2.609)
Sh. migr. x Sh. CBA cov. x Q3 sh. skilled		-2.862		-2.802
		(2.935)		(2.821)
Sh. migr. x Sh. CBA cov. x Q4 sh. skilled		-5.361		-5.113
		(3.096)		(3.021)
Sh. migr. x Sh. CBA cov. x Q5 sh. skilled		-5.857		-5.709
		(3.769)		(3.740)
N	636	636	636	636

Note: Table presents estimates from IV regressions using commuting zone-level data. Share of migrants is the number of foreign residents in a commuting zone in a year divided by native population in a commuting zone in 2000. Share skilled is the share of native residents with upper-secondary or higher level of education in 2000. Controls are listed in Table A2; all specifications include commuting zone fixed effects. Weights assigned to observations equal the number of Swiss residents in 2000. Standard errors in parentheses are clustered at the commuting zone level. Source: FSO, SECO, ZEMIS.

Table A7: Wage analysis by percentiles of the wage distribution: IV robustness checks

	Outcome: ln **real gross hourly wage at the m-th percentile**			
	50th pct	5th pct	10th pct	95th pct
	(1)	(2)	(3)	(4)
Panel A: 1970				
Sh. migrants	-0.109	-0.620	-0.742	0.278
	(0.162)	(0.309)	(0.327)	(0.275)
Mean outcome	3.553	3.062	3.157	4.224
Sd outcome	0.103	0.070	0.068	0.195
First stage F-stat	6.243	6.243	6.243	6.243
N	848	848	848	848
Panel B: Interaction 4-digit industry				
Sh. migrants	-0.177	-1.493	-1.531	0.671
	(0.482)	(0.527)	(0.591)	(0.658)
Sh. migr. x Sh. CBA cov.	0.147	5.626	5.375	-1.415
	(2.189)	(2.228)	(2.440)	(2.775)
Mean outcome	3.553	3.062	3.157	4.224
Sd outcome	0.103	0.070	0.068	0.195
First stage F-stat	4.300	4.300	4.300	4.300
N	848	848	848	848

Note: The table presents estimates from IV regressions using biennial data at the commuting zone level between 2000 and 2014. Share of migrants is the number of foreign residents in a commuting zone in a year divided by native population in a commuting zone in 2000. Controls are listed in Table A2; all specifications include commuting zone and year fixed effects. Weights assigned to observations equal the number employed in commuting zone in 2000. Standard errors in parentheses are clustered at the commuting zone level. Source: FSO, SECO, SESS, ZEMIS.

Table A8: Wage analysis by native educational attainment and CBA coverage

	Outcome: mean ln **of native gross hourly wage**			
	All	Up to lower-secondary	Upper-secondary	Tertiary
	(1)	(2)	(3)	(4)
Panel A: IV with controls				
Sh. migrants	-0.107	-1.040	-0.347	0.082
	(0.123)	(0.452)	(0.108)	(0.109)
Mean outcome	3.594	3.345	3.526	3.879
Sd outcome	0.109	0.082	0.081	0.103
First stage F-stat	9.722	9.722	9.722	9.721
N	848	848	848	847
Panel B: IV interaction with controls				
Sh. migrants	-0.207	-1.944	-0.577	0.089
	(0.384)	(0.667)	(0.318)	(0.265)
Sh. migr. x Sh. CBA cov.	0.449	5.332	1.422	-0.064
	(1.530)	(2.298)	(1.273)	(1.165)
Mean outcome	3.594	3.345	3.526	3.879
Sd outcome	0.109	0.082	0.081	0.103
First stage F-stat	3.638	3.638	3.638	3.638
N	848	848	848	847

Note: The table presents estimates from IV regressions using biennial data at the commuting zone level between 2000 and 2014. Share of migrants is the number of foreign residents in a commuting zone in a year divided by native population in a commuting zone in 2000. Controls are listed in Table A2; all specifications include commuting zone and year fixed effects. Lower-secondary level of education is compulsory education as highest degree, upper-secondary is an apprenticeship or a matura, tertiary is a degree from a university, university of applied sciences, university of teacher education or a professional degree. Weights assigned to observations equal the number of natives employed in commuting zone in 2000. Standard errors in parentheses are clustered at the commuting zone level. Source: FSO, SECO, SESS, ZEMIS.

Table A9: Employment analysis by educational level: IV robustness checks

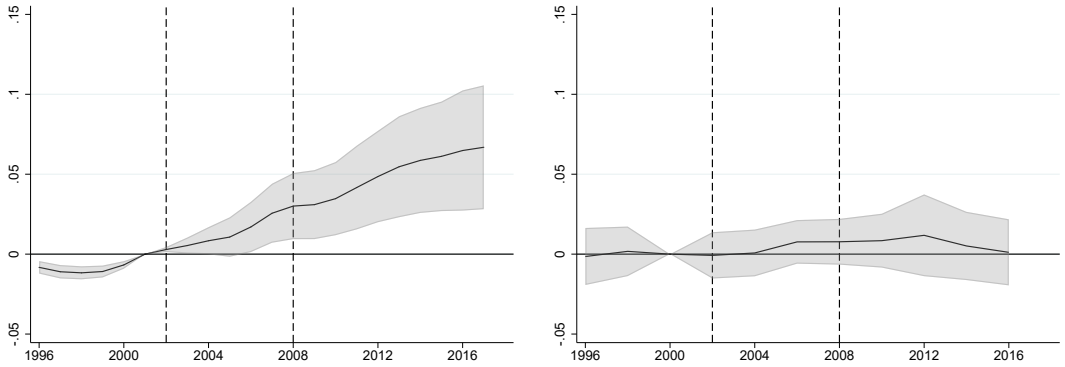
	Outcome: share of natives employed in population 18-65			
	All	Up to lower-secondary	Upper-secondary	Tertiary
	(1)	(2)	(3)	(4)
Panel A: 1970				
Sh. migrants	-0.658	0.525	-1.021	-0.644
	(0.401)	(0.484)	(0.443)	(0.373)
Mean outcome	0.776	0.451	0.787	0.909
Sd outcome	0.047	0.117	0.059	0.053
First stage F-stat	5.562	5.548	5.562	5.557
N	1590	1576	1590	1585
Panel B: Interaction 4-digit industry				
Sh. migrants	-1.044	-0.865	-1.118	-0.739
	(0.584)	(0.933)	(0.546)	(0.608)
Sh. migr. x Sh. CBA cov.	3.902	3.639	3.715	2.251
	(2.325)	(4.209)	(2.164)	(2.598)
Mean outcome	0.776	0.451	0.787	0.909
Sd outcome	0.047	0.117	0.059	0.053
First stage F-stat	3.467	3.444	3.467	3.464
N	1590	1576	1590	1585

Note: The table presents estimates from IV regressions using annual data at the commuting zone level between 2000 and 2014. Share of migrants is the number of foreign residents in a commuting zone in a year divided by native population in a commuting zone in 2000. Controls are listed in Table A2; all specifications include commuting zone and year fixed effects. Lower-secondary level of education is compulsory education as highest degree, upper-secondary is an apprenticeship or a matura, tertiary is a degree from a university, university of applied sciences, university of teacher education or a professional degree. Weights assigned to observations equal the number of native respondents 18-65 years of age in commuting zone in 2000. Standard errors in parentheses are clustered at the commuting zone level. Source: FSO, SECO, SLFS, ZEMIS.

Appendix B (for Chapter 2)

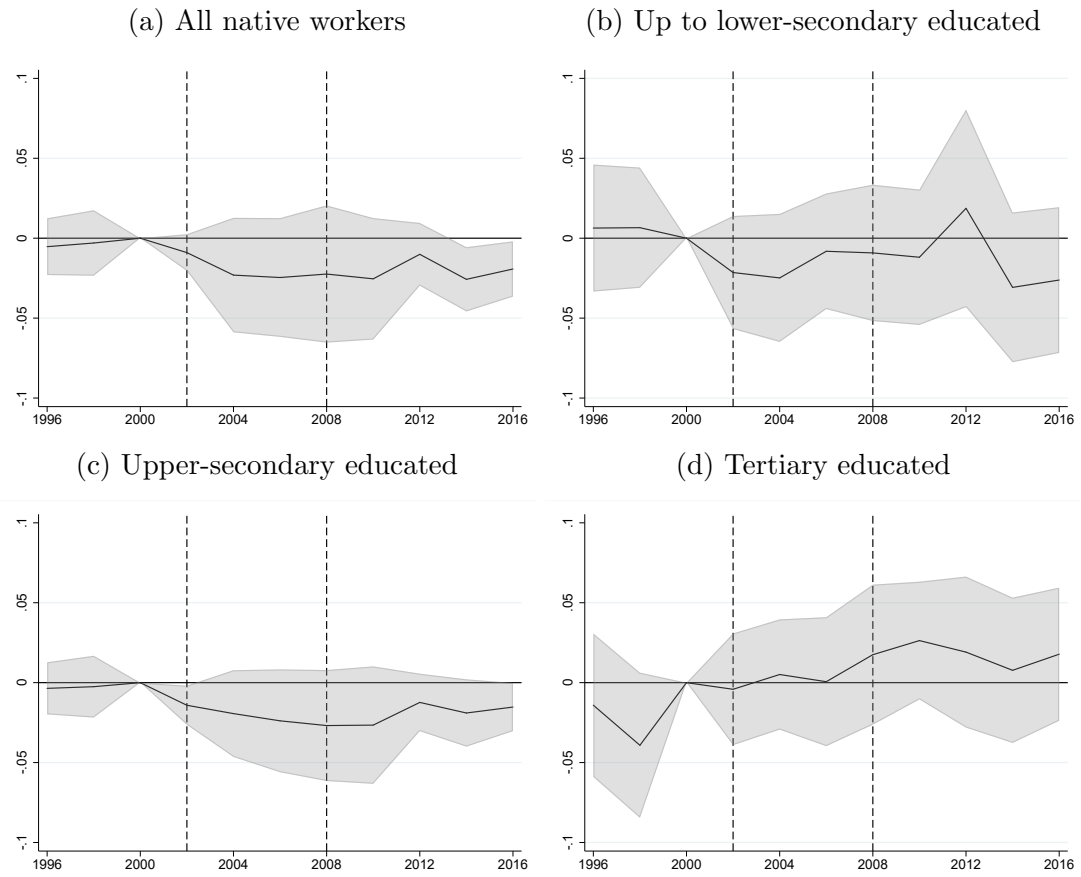
Figure B1: Exposure to cross-border commuters and foreign workers

(a) Cross-border commuters (b) Resident migrant workers



Note: This figure shows difference-in-differences estimates using annual (biennial) data at the commuting zone level for the period 1996–2017 (1996–2016) in Panel a (b). The vertical lines indicate the beginning of the transition period (2002) and the beginning of the post-treatment period (2008). Outcome is the number of cross-border commuters divided by total employment in 1995 in Panel (a) and number of resident foreign workers (excluding cross-border commuters) divided by total employment in 1996 in Panel (b). Weights assigned to observations equal total employment in first year. Standard errors are clustered at the commuting zone level, 95% confidence intervals shown. Sources: FSO in Panel (a) and SESS in Panel (b).

Figure B2: Wages by educational level



Note: This figure shows difference-in-differences estimates using biennial data at the commuting zone level for the period 1996–2016. The vertical lines indicate the beginning of the transition period (2002) and of the post-treatment period (2008). Outcome is the mean natural log of gross hourly wage of natives. Observations are weighed by number of native employees in a specific education category in first year. Standard errors are clustered at the commuting zone level, 95% confidence intervals shown. Source: SESS.

Table B1: Exposure to cross-border commuters by educational level (robustness checks to treatment definition)

	Outcome: share of cross-border commuters			
	All	Up to lower-secondary	Upper-secondary	Tertiary
	(1)	(2)	(3)	(4)
Panel A: Threshold value				
25min * 2002-2007	0.015	-0.002	0.011	0.006
	(0.007)	(0.002)	(0.005)	(0.002)
25min * after 2008	0.067	0.007	0.038	0.023
	(0.019)	(0.004)	(0.011)	(0.007)
35min * 2002-2007	0.012	-0.002	0.010	0.004
	(0.005)	(0.002)	(0.004)	(0.002)
35min * after 2008	0.050	0.004	0.031	0.016
	(0.015)	(0.003)	(0.008)	(0.005)
Mean outcome	0.070	0.020	0.038	0.012
Sd outcome	0.115	0.053	0.056	0.021
Commuting zones	106	106	106	106
within 25 min	28	28	28	28
within 35 min	41	41	41	41
N	1166	1166	1166	1166
Panel B: Continuous treatment				
Travel time * 2002-2007	0.017	-0.003	0.013	0.007
	(0.008)	(0.002)	(0.006)	(0.003)
Travel time * after 2008	0.079	0.007	0.045	0.027
	(0.022)	(0.005)	(0.013)	(0.008)
Mean outcome	0.070	0.020	0.038	0.012
Sd outcome	0.115	0.053	0.056	0.021
Commuting zones	106	106	106	106
N	1166	1166	1166	1166

Note: This table shows difference-in-differences estimates using biennial data at the commuting zone level for the period 1996–2016. The continuous measure applies the function $\exp(-ax)$ with $a = 0.05$ to travel time. Outcome is the share of cross-border commuters in total employed. Denominator is fixed at first year. Observations are weighed by total workforce in first year. Standard errors in parentheses are clustered at the commuting zone level. Sources: SESS.

Table B2: Enrolment by institutional type (robustness checks to treatment definition)

	Outcome: share of enrolled native first-year students			
	All	University	University of applied sciences	University of teacher education
	(1)	(2)	(3)	(4)
Panel A: Threshold value				
25min * 2002-2007	0.005	-0.001	0.006	0.001
	(0.008)	(0.006)	(0.004)	(0.004)
25min * after 2008	0.023	0.003	0.022	-0.002
	(0.011)	(0.009)	(0.005)	(0.004)
35min * 2002-2007	0.010	0.002	0.008	0.001
	(0.006)	(0.005)	(0.003)	(0.003)
35min * after 2008	0.038	0.015	0.022	0.000
	(0.010)	(0.008)	(0.005)	(0.004)
Mean outcome	0.372	0.207	0.136	0.035
Sd outcome	0.119	0.086	0.052	0.020
Commuting zones	106	106	106	106
within 25 min	28	28	28	28
within 35 min	41	41	41	41
N	2226	2226	2226	1802
Panel B: Continuous treatment				
Travel time * 2002-2007	0.007	-0.000	0.009	-0.000
	(0.010)	(0.007)	(0.005)	(0.004)
Travel time * after 2008	0.028	0.007	0.025	-0.004
	(0.013)	(0.011)	(0.006)	(0.005)
Mean outcome	0.372	0.207	0.136	0.035
Sd outcome	0.119	0.086	0.052	0.020
Commuting zones	106	106	106	106
N	2226	2226	2226	1802
Panel C: Municipality level				
30min * 2002-2007	0.013	0.006	0.007	0.002
	(0.005)	(0.004)	(0.002)	(0.002)
30min * after 2008	0.035	0.014	0.021	0.001
	(0.006)	(0.005)	(0.003)	(0.002)
Mean outcome	0.372	0.206	0.136	0.035
Sd outcome	0.181	0.131	0.090	0.042
Municipalities	2222	2222	2222	2222
within 30 min	785	785	785	785
N	46662	46662	46662	37774

Note: This table shows difference-in-differences estimates using annual data at the commuting zone level (municipality level in Panel C) for the period 1997–2017. The continuous measure applies the function $\exp(-ax)$ with $a = 0.05$ to travel time. Outcome is the share of native first-year students in birth cohort. Denominator is fixed at first year and specific to the education category. Observations are weighed by cohort size in a specific education category in first year. Standard errors in parentheses are clustered at the commuting zone level (municipality level in Panel C). Source: SHIS-studex.

Table B3: Enrolment at UAS (robustness checks)

	Outcome: share of enrolled native first-year students			
	Baseline	+ Education supply	+ Labor demand	No weights
	(1)	(2)	(3)	(4)
30min * 2002-2007	0.003	0.003	0.003	0.002
	(0.004)	(0.004)	(0.004)	(0.004)
30min * after 2008	0.018	0.018	0.017	0.015
	(0.005)	(0.005)	(0.005)	(0.005)
Bartik control			0.038	
			(0.047)	
Mean outcome	0.136	0.136	0.136	0.130
Sd outcome	0.052	0.052	0.052	0.054
Commuting zones	106	106	106	106
within 30 min	35	35	35	35
N	2226	2226	2226	2226

Note: This table shows difference-in-differences estimates using annual data at the commuting zone level for the period 1997–2017. Outcome is the share of native first-year students in birth cohort. Denominator is fixed at first year and specific to the education category. Observations are weighed by cohort size in a specific education category in first year. Column (1) is the baseline specification from Table 2.3, columns (2) and (3) include additional control variables. We use two education supply controls – a dummy variable for an institution and the number of study fields at the ISCED level available within a 20km radius of the main city of the region. The Bartik control is based on employment data. Column (4) is unweighed. Standard errors in parentheses are clustered at the commuting zone level. Source: SHIS-studex.

Table B4: Employment rates by educational level

	Outcome: native employment rate			
	All	Up to lower-secondary	Upper-secondary	Tertiary
	(1)	(2)	(3)	(4)
30min * 2002-2007	0.008	0.033	0.006	-0.006
	(0.008)	(0.026)	(0.010)	(0.009)
30min * after 2008	-0.003	-0.017	0.004	-0.003
	(0.008)	(0.022)	(0.010)	(0.011)
Mean outcome	0.777	0.458	0.788	0.907
Sd outcome	0.046	0.106	0.056	0.048
Commuting zones	106	106	106	104
within 30 min	35	35	35	35
N	1166	1162	1166	1143

Note: This table shows difference-in-differences estimates using biennial data at the commuting zone level for the period 1996–2016. Outcome is the number of native employed divided by the total number of respondents. Denominator is fixed at first year and specific to the education category. Observations are weighed by the number of respondents in a specific education category in the first year. Standard errors in parentheses are clustered at the commuting zone level. Source: SLFS.

Table B5: Graduation rates of first-year students by institutional type

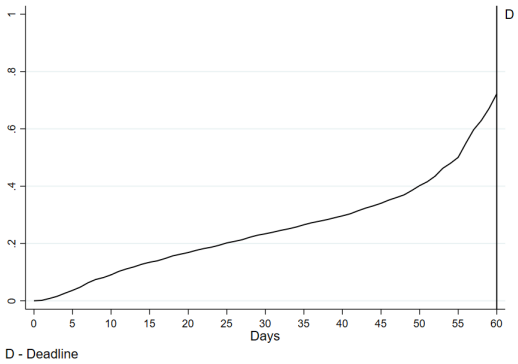
	Outcome: graduation rate			
	All	University	University of applied sciences	University of teacher education
	(1)	(2)	(3)	(4)
30min * 2002-2007	0.004	0.002	0.008	-0.106
	(0.005)	(0.008)	(0.009)	(0.051)
30min * after 2008	0.003	0.001	0.006	-0.092
	(0.006)	(0.008)	(0.009)	(0.061)
Mean outcome	0.693	0.633	0.676	0.687
Sd outcome	0.305	0.289	0.295	0.353
Commuting zones	106	106	106	106
within 30 min	35	35	35	35
N	2226	2226	2224	1749

Note: This table shows difference-in-differences estimates using annual data at the commuting zone level for the period 1997–2017. Outcome is the share of native first-year students that graduated within 1997–2017 relative to the number enrolled. Observations are weighed by cohort size in a specific education category in first year. Standard errors in parentheses are clustered at the commuting zone level. Source: SHIS-studex.

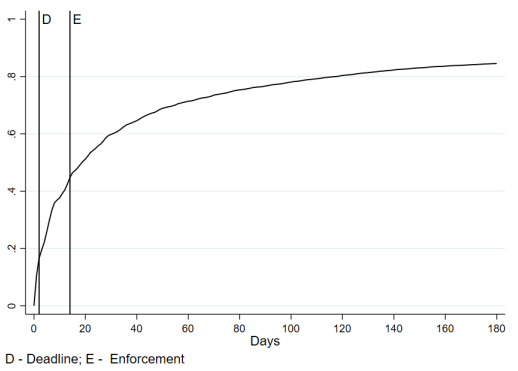
Appendix C (for Chapter 3)

Figure C1: Dynamics of Tax Compliance in the Control Group

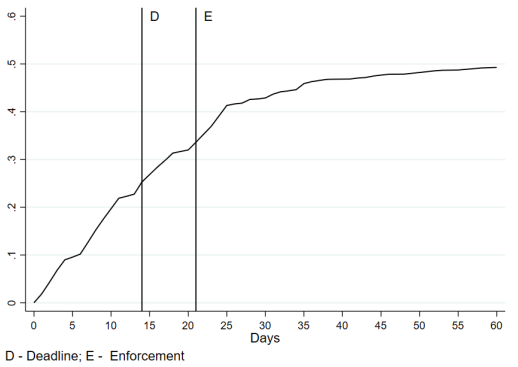
(a) Tax Payment



(b) Payment Reminder



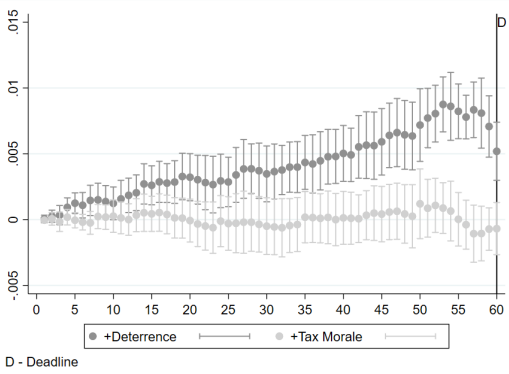
(c) Filing Reminder



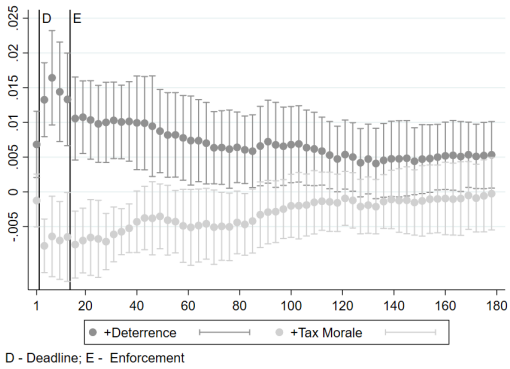
Note: The figure presents average compliance in the control group by days since letter receipt for the TP (Panel (a)), TPR FY2014 (Panel (b)) and TFR FY2015 (Panel (c)) experiments. Outcome is partial payment probability at 60 days / deadline in Figure (a) and at 14 days / enforcement start in Figure (b); outcome is filing probability at 21 days / enforcement start in Figure (c).

Figure C2: Dynamic Effects of Deterrence and Tax Morale Messages

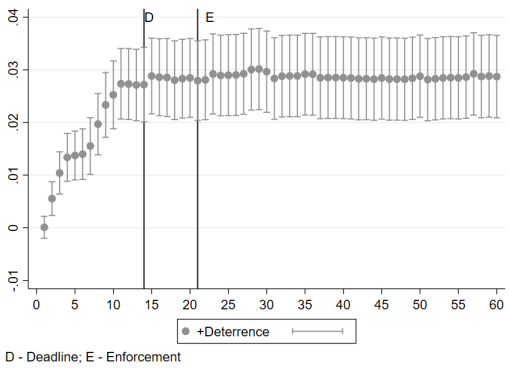
(a) Tax Payment



(b) Payment Reminder



(c) Filing Reminder

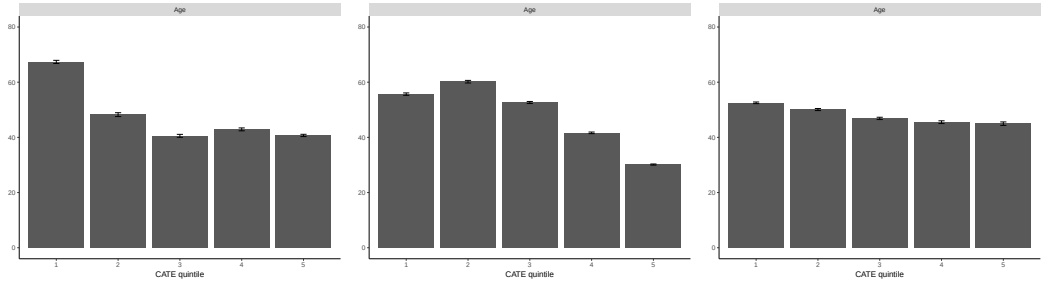


Note: The figure presents deterrence and tax morale treatment effect estimates by days since letter receipt for the TP (Panel (a)), TPR FY2014 (Panel (b)) and TFR FY2015 (Panel (c)) experiments. The outcome is partial payment probability in Panels (a) and (b), and filing probability in Panel (c). The vertical lines indicate the payment/filing deadline and/or the day follow-up enforcement starts. Controls are listed in Table 3.1. 95% confidence intervals based on robust standard errors are plotted. Standard errors are clustered by date of letter receipt in Panels (a) and (b).

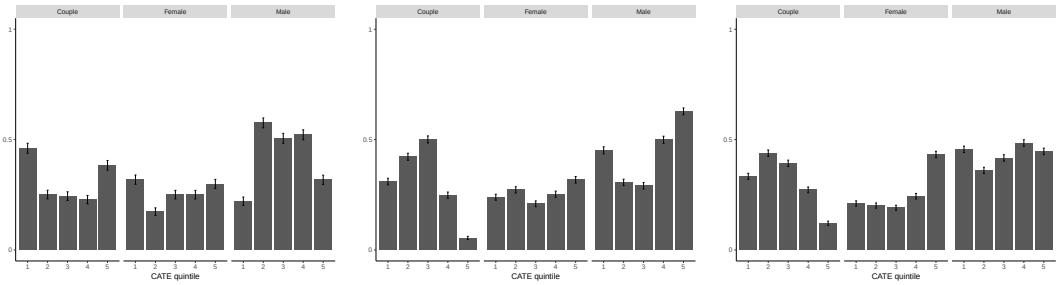
Figure C3: Average Value of Control Variables by Quintile of Treatment Effects

Simplification **Deterrence** **Tax Morale**

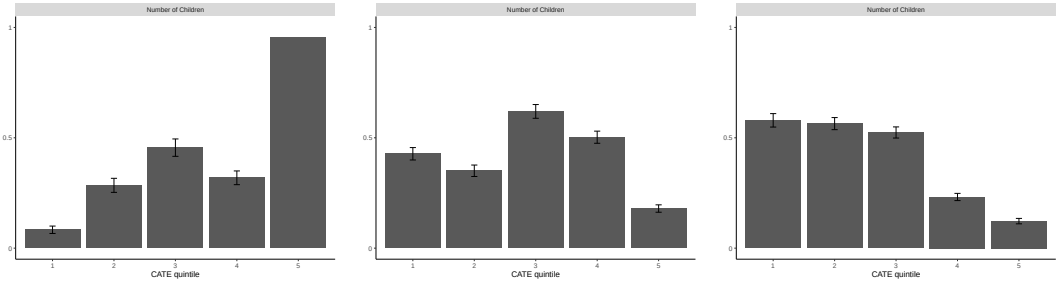
(a) Average Age



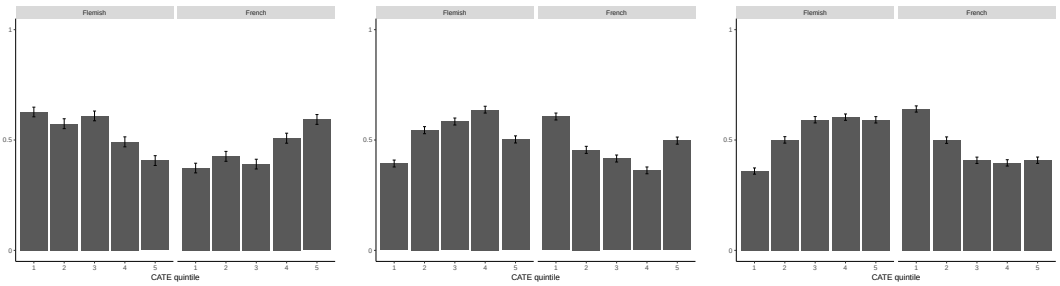
(b) Average of Gender categories



(c) Average Number of Children

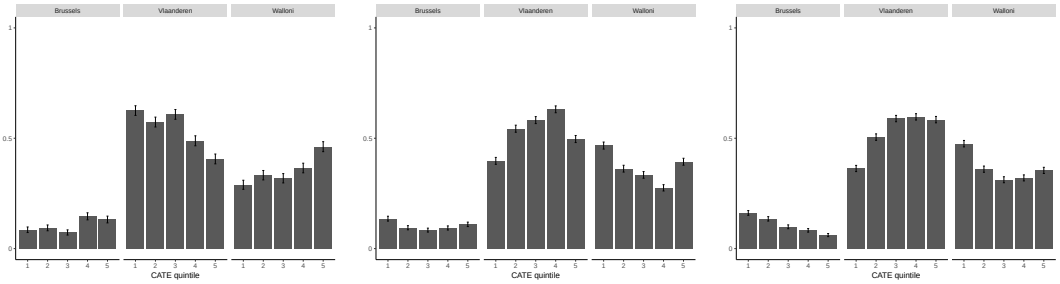


(d) Average of Language Categories

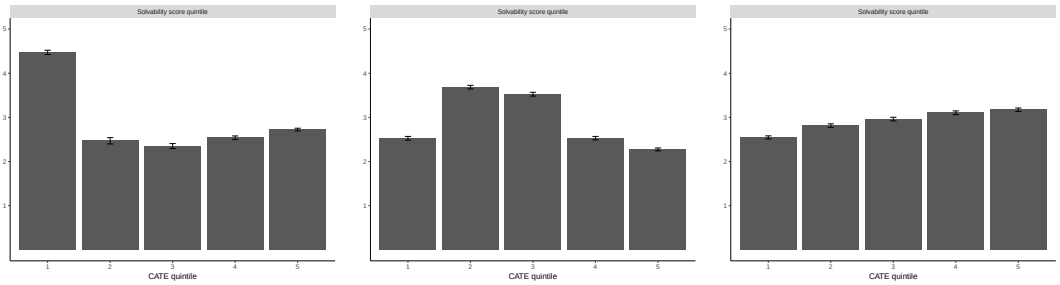


Simplification **Deterrence** **Tax Morale**

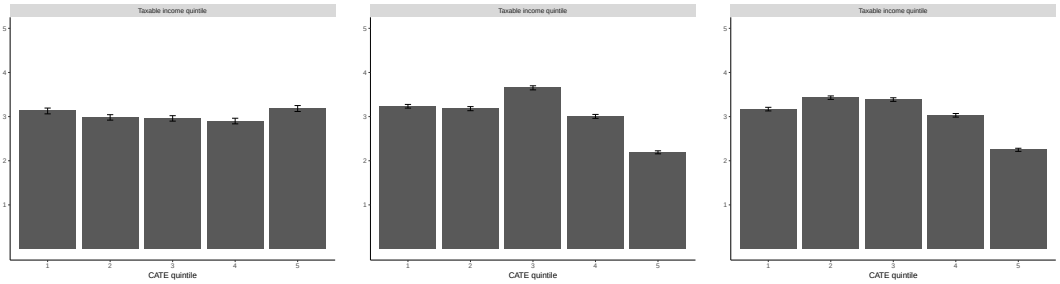
(e) Average of Region Categories



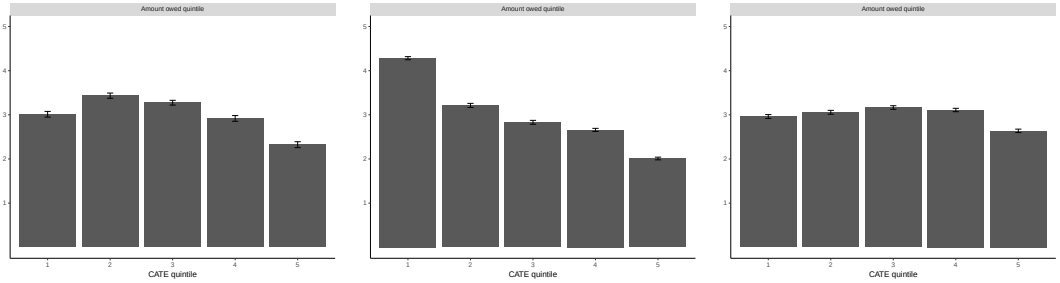
(f) Average Solvency Score



(g) Average Taxable Income



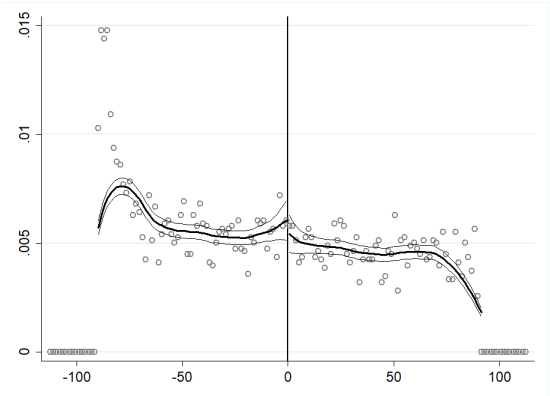
(h) Average Tax Liability



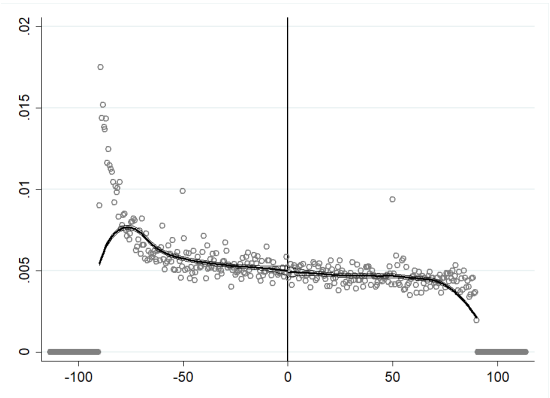
Note: The figure presents the mean and 95% confidence interval of control variables from TPR FY2014 experiment by quintile of conditional average treatment effect (CATE). These were estimated using the generalized random forest (GRF) algorithm (Wager and Athey, 2018). Three panels in each figure differ in the definition of treatment and control groups. The underlying sample of taxpayers are those in the control group and those sent a simplified letter without additional messages in the left panel, simplified letter and a simplified letter with a deterrence message in the middle panel, a simplified letter and a simplified letter with a tax morale message in right panel.

Figure C4: RDD – Identifying Assumptions

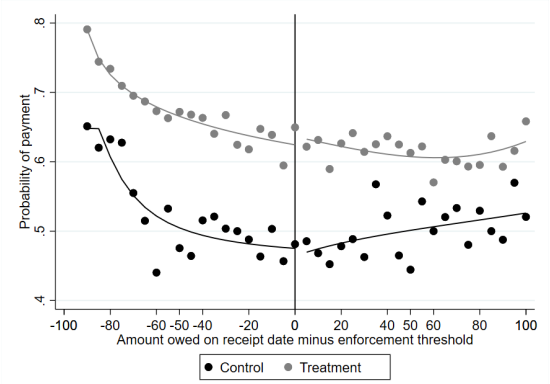
(a) Density around the threshold - Control



(b) Density around the threshold - Treatment



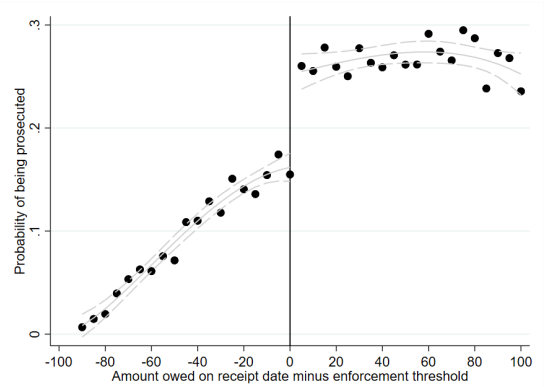
(c) Probability of Paying before Enforcement



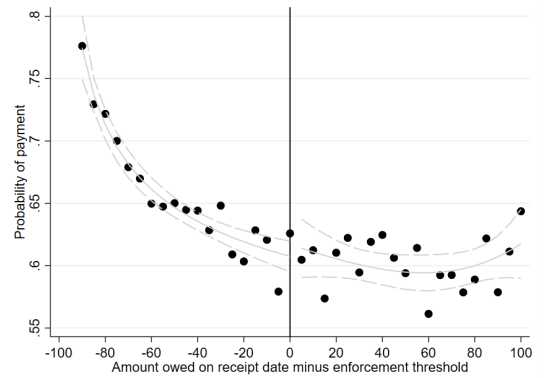
Note: The figure is based on the TPR FY2014 experiment. It explores the plausibility of the identification assumptions underlying the RDD. Panels (a) and (b) plot the average density by bin in the control and treatment group, respectively. Panel (c) plots the probability of payment before enforcement by initial amount owed (centred at the enforcement threshold). Bin size is set to €5 and amounts within €100 of the enforcement threshold are considered. Fractional polynomial predictions are plotted as well.

Figure C5: Effects of Enforcement

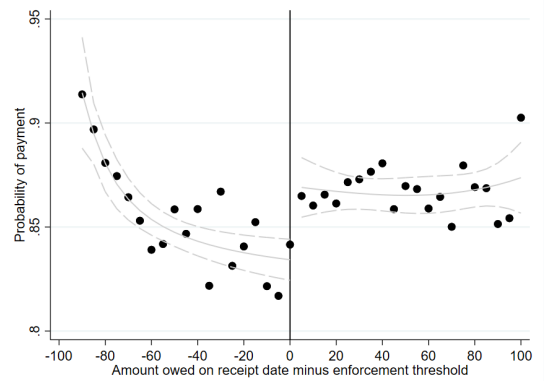
(a) Probability of Enforcement at 180 days



(b) Probability of Partial Payment at 14 days / enforcement start



(c) Probability of Partial Payment at 180 days



Note: The figure is based on the the TPR FY2014 experiment. It shows probability of enforcement after 180 days (Panel (a)), probability of paying after 14 days (Panel (b)) and probability of paying after 180 days (Panel (c)) by initial amount owed (centred at the enforcement threshold). Bin size is set to €5 and amounts within €100 of the enforcement threshold are considered. Fractional polynomial predictions with 95% confidence intervals plotted.

Table C1: Deterrence and Tax Morale Messages by Experiment

Experiment / Type	Name	Message
Panel A: Tax Payment		
Deterrence	Explicit Penalty	These costs amount to 209 euros on average and can go up depending on the circumstances.
	Enforcement + Immediacy	Warning: do not wait until the deadline to pay, you run the risk of being late. If you do not pay on time, we will start actions to recover this amount.
Tax Morale	Social Norm	In Belgium 95% of taxes are paid on time.
	Public Goods	Tax revenues allow basic public services such as health care, education and law and order, to function.
Panel B: Payment Reminders		
Deterrence	Explicit Penalty (EP) (FY2014, 2015)	These costs amount to 209.00 euro on average and may, depending on the situation, rise further.
	Active Choice (FY2014)	Not paying your taxes will be seen as an active choice.
	EP + Immediacy (FY2015)	These costs amount to 209.00 euro on average and may, depending on the situation, rise further. By paying now you may still avoid these costs.
	EP + Enforcement (FY2015)	These costs amount to 209.00 euro on average and may, depending on the situation, rise further. We will undertake actions to claim tax dues that may involve seizing your income or your assets.
	EP (Female Name First) (FY2015)	Woman's name, Man's name (instead of reversed)
Tax Morale	Social Norm (FY2014, 2015)	You belong to a minority of taxpayers who did not pay their taxes within the legal period: 95% of taxes in Belgium are paid on time. Why not follow this example?
	Public Goods (FY2014)	Paying taxes guarantees the provision of essential services by the government, such as public health, education, and public safety.
	Public Goods Negative (FY2014, 2015)	Not paying taxes puts at risk the provision of essential services by the government, such as public health, education, and public safety.
Panel C: Tax Filing		
Tax Morale	Public Goods	The above pie chart illustrates how your taxes and social security contributions are spent in terms of public services.
	Public Goods Negative	The above pie chart illustrates how your taxes and social security contributions are spent in terms of public services. Incorrect and timely completion of the declaration puts at risk the essential services provided by the government.
	Public Goods + Penalty	The above pie chart illustrates how your taxes and social security contributions are spent in terms of public services. By completing your declaration correctly and in a timely fashion, you avoid further measures such as fines and tax increases.
	Public Goods + Social Norms	The above pie chart illustrates how your taxes and social security contributions are spent in terms of public services. The vast majority of people complete their declaration correctly and in a timely manner. Please follow this example.
Panel D: Filing Reminders		
Deterrence	Explicit Penalty	You risk a penalty of 50 to 1,250 euro and a tax increase of 10 to 200%.

Note: The table lists all letter messages by experiment and treatment type. In all experiments the messages where added to a personalized simplified letter.

Table C2: Randomization Design for TPR, TF and TFR experiments (using national identity number)

Digits	TPR2014	TFR2015	TPR2015	TF2016	Digits	TPR2014	TFR2015	TPR2015	TF2016
01	C	C	C	S+PG	**51**	S+PG	S+EP	S+PGN+EP	C
02	C	C	S	S+PG	**52**	S+PG	S+EP	S+SN	C
03	C	C	S+EP	S+PG	**53**	S+PG	S+EP	S+SN+EP	C
04	C	C	S+EP FM	S+PG	**54**	S+SN+PG	S+EP	C	C
05	C	C	S+EP ENF	S+PG	**55**	S+SN+PG	S+EP	S	C
06	C	C	S+EP IMM	S+PG	**56**	S+SN+PG	S+EP	S+EP	C
07	C	C	S+PGN	S+PG	**57**	S+SN+PG	S+EP	S+EP FM	C
08	C	C	S+PGN+EP	S+PG	**58**	S+SN+PG	S+EP	S+EP ENF	C
09	C	C	S+SN	S+PG	**59**	S+SN+PG	S+EP	S+EP IMM	C
10	C	C	S+SN+EP	S+PG	**60**	S+SN+PG	S+EP	S+PGN	C
11	C	S	C	S+PG	**61**	S+SN+PG	S+EP	S+PGN+EP	C
12	S	S	C	S+PG	**62**	S+SN+PG	S+EP	S+SN	C
13	S	S	S	S+PGN	**63**	S+SN+PG	S+EP	S+SN+EP	C
14	S	S	S+EP	S+PGN	**64**	S+SN+PG	S+EP	S+PGN	C
15	S	S	S+EP FM	S+PGN	**65**	S+AC	S+EP	C	C
16	S	S	S+EP ENF	S+PGN	**66**	S+AC	S+EP	S	C
17	S	S	S+EP IMM	S+PGN	**67**	S+AC	S+EP	S+EP	C
18	S	S	S+PGN	S+PGN	**68**	S+AC	S+EP	S+EP FM	C
19	S	S	S+PGN+EP	S+PGN	**69**	S+AC	S+EP	S+EP ENF	C
20	S	S	S+SN	S+PGN	**70**	S+AC	S+EP	S+EP IMM	C
21	S	S	S+SN+EP	S+PGN	**71**	S+AC	S+EP	S+PGN	C
22	S	S+EP	S	S+PGN	**72**	S+AC	S+EP	S+PGN+EP	C
23	S+SN	S+EP	C	S+PGN	**73**	S+AC	S+EP	S+SN	C
24	S+SN	S+EP	S	S+PGN	**74**	S+AC	S+EP	S+SN+EP	C
25	S+SN	S+EP	S+EP	S+PG+SN	**75**	S+AC	S+EP	S+EP	C
26	S+SN	S+EP	S+EP FM	S+PG+SN	**76**	S+EP	S+EP	C	C
27	S+SN	S+EP	S+EP ENF	S+PG+SN	**77**	S+EP	S+EP	S	C
28	S+SN	S+EP	S+EP IMM	S+PG+SN	**78**	S+EP	S+EP	S+EP	C
29	S+SN	S+EP	S+PGN	S+PG+SN	**79**	S+EP	S+EP	S+EP FM	C
30	S+SN	S+EP	S+PGN+EP	S+PG+SN	**80**	S+EP	S+EP	S+EP ENF	C
31	S+SN	S+EP	S+SN	S+PG+SN	**81**	S+EP	S+EP	S+EP IMM	C
32	S+SN	S+EP	S+SN+EP	S+PG+SN	**82**	S+EP	S+EP	S+PGN	C
33	S+SN	S+EP	S+SN	S+PG+SN	**83**	S+EP	S+EP	S+PGN+EP	C
34	S+PGN	S+EP	C	S+PG+SN	**84**	S+EP	S+EP	S+SN	C
35	S+PGN	S+EP	S	S+PG+SN	**85**	S+EP	S+EP	S+SN+EP	C
36	S+PGN	S+EP	S+EP	S+PG+SN	**86**	S+AC+EP	S+EP	S+EP	C
37	S+PGN	S+EP	S+EP FM	S+PG+EP	**87**	S+AC+EP	S+EP	C	C
38	S+PGN	S+EP	S+EP ENF	S+PG+EP	**88**	S+AC+EP	S+EP	S	C
39	S+PGN	S+EP	S+EP IMM	S+PG+EP	**89**	S+AC+EP	S+EP	S+EP	C
40	S+PGN	S+EP	S+PGN	S+PG+EP	**90**	S+AC+EP	S+EP	S+EP FM	C

TREATMENT ALLOCATION

C	**Control**
S	**Simplification**
Deterrence Messages	
EP	Explicit Penalty (EP)
AC	Active Choice (AC)
EP FM	EP (Female Name First)
EP ENF	EP+Enforcement
EP IMM	EP+Immediacy
Tax Morale Treatments	
PG	Public Goods (PG)
PGN	PG Negative (PGN)
SN	Social Norms (SN)

41	S+PGN	S+EP	S+PGN+EP	S+PG+EP	**91**	S+AC+EP	S+EP	S+EP ENF	C
42	S+PGN	S+EP	S+SN	S+PG+EP	**92**	S+AC+EP	S+EP	S+EP IMM	C
43	S+PGN	S+EP	S+SN+EP	S+PG+EP	**93**	S+AC+EP	S+EP	S+PGN	C
44	S+PG	S+EP	C	S+PG+EP	**94**	S+AC+EP	S+EP	S+PGN+EP	C
45	S+PG	S+EP	S	S+PG+EP	**95**	S+AC+EP	S+EP	S+SN	C
46	S+PG	S+EP	S+EP	S+PG+EP	**96**	S+AC+EP	S+EP	S+SN+EP	C
47	S+PG	S+EP	S+EP FM	S+PG+EP	**97**	S+AC+EP	S+EP	S+EP	C
48	S+PG	S+EP	S+EP ENF	S+PG+EP					
49	S+PG	S+EP	S+EP IMM	C					
50	S+PG	S+EP	S+PGN	C					

Note: The table presents the randomization design of four separate experiments. TPR stands for payment reminders, TFR for filing reminders and TF to tax filing experiment. 2-digits are the last two digits of the national identity number. See Appendix Table C1 for more details on treatment messages. All messages were added to personalized simplified letters.

Table C3: Randomization Design for TP experiment (using Day of Birth)

Day	TP	Day	TP
01	C	**17**	S + PG
02	C	**18**	S + PG
03	C	**19**	S + PG
04	C	**20**	S + PG
05	S (NP)	**21**	S + ENF+ IMM
06	S (NP)	**22**	S + ENF+ IMM
07	S (NP)	**23**	S + ENF+ IMM
08	S (NP)	**24**	S + ENF+ IMM
09	S + EP	**25**	S
10	S + EP	**26**	S
11	S + EP	**27**	S
12	S + EP	**28**	S
13	S + SN	**29**	S
14	S + SN	**30**	S
15	S + SN	**31**	S
16	S + SN		

Treatment Groups

C	**Control**
S	**Simplification**
S (NP)	**Simplification (Not Pers.)**
Deterrence Messages	
S + EP	Explicit Penalty
S + ENF+ IMM	Enforcement+Immediacy
Tax Morale Messages	
S + PG	Public Goods
S + SN	Social Norms

Note: The table presents the randomization design of the Tax Payment (TP) experiment. Day stands for the day the taxpayer was born. Simplified letters that were not personalized started with "Mr., Ms." instead of "Mr X" or "Ms X" (where X is the name of the taxpayer). See Appendix Table C1 for more details on treatment messages. All messages were added to personalized simplified letters.

Table C4: Overlap across experiments

	Share of taxpayers in experiment		
	Payment Reminders FY2014	Payment Reminders FY2015	Filing Reminders FY2015
Experiment	(1)	(2)	(3)
Payment Reminders FY2014	1.000	0.283	0.062
Payment Reminders FY2015	0.307	1.000	0.066
Filing Reminders FY2015	0.106	0.104	1.000

Note: The table presents the overlap between populations of taxpayers in the payment reminders (TPR) and filing reminders experiments (TFR). Each cell gives the share of taxpayers in the experiment listed horizontally that were also part of the population of the experiment listed vertically.

Table C5: Filing Reminders FY2015 controlling for TPR FY2014 treatment assignment

	Probability of having filed 21 days (before enforcement) (1)
Simplified (S)	0.019
	(0.011)
+ Deterrence	0.029
	(0.010)
P-values of tests:	
Simplified=Control	0.072
S+Deterrence=Simplified	0.005
Control mean	0.317
N	148,925

Note: The table presents treatment effect estimates from filing reminders experiment (TFR FY2015). Control variables are listed in Table 3.1. Additional controls include dummies for the treatment the taxpayer would have received if had been late with payment in the previous fiscal year. Robust standard errors in parentheses.

Table C6: Payment Experiments: Individual Letter Effects

Probability of some payment	at 14 days (before enforcement)		at 60 days (deadline)
	TPR FY2014 (1)	TPR FY2015 (2)	TP (3)
Simplification Treatments			
Simplified (S)	0.102	0.107	0.005
	(0.010)	(0.004)	(0.002)
+ Not Personalized (NP)			0.001
			(0.002)
Deterrence Treatments			
+ Explicit Penalty (EP)	0.020	0.009	0.004
	(0.002)	(0.003)	(0.001)
+ Active Choice (AC)	0.001		
	(0.004)		
+ EP + AC	0.016		
	(0.005)		
+ EP + Enforcement		0.024	
		(0.003)	
+ EP + Immediacy		0.017	
		(0.004)	
+ EP FM		0.008	
		(0.005)	
+ Enforcement + Immediacy			0.007
			(0.001)
Tax Morale Treatments			
+ Public Goods Negative (PGN)	-0.007	-0.014	
	(0.004)	(0.003)	
+ Public Goods Positive (PGP)	-0.014		-0.002
	(0.004)		(0.001)
+ Social Norms (SN)	-0.002	-0.011	0.001
	(0.004)	(0.004)	(0.001)
+ SN + PGP	-0.006		
	(0.004)		
Deterrence & Tax Morale Treatments			
+ EP + SN		0.006	
		(0.003)	
+ EP + PGN		0.005	
		(0.005)	
P-values of tests:			
Simplified=Control	0.001	0.001	0.001
S + NP = S			0.498
S + EP = S	0.001	0.451	0.017
S + AC = S	0.859		
S + EP + AC = S + EP	0.491		
S + EP + Enforcement = S + EP		0.001	
S + EP + Immediacy = S + EP		0.077	
S + EP FM = S + EP		0.916	
S + Enforcement + Immediacy = S			0.001
S + PGN = S	0.61	0.001	
S + PGP = S	0.007		0.262
S + SN = S	0.92	0.375	0.651
S + SN + PGP = S	0.562		
S + EP + SN = S + EP		0.956	
S + EP + PGN = S + EP		0.991	
Control mean	0.447	0.418	0.728
N	229,751	202,730	1,216,317

Note: The table presents treatment effect estimates of messages in the two payment reminder experiments (TPR 2014 in column 1 and TPR 2015 in column 2) and in the tax payment (TP) experiment (column 3). Control variables are listed in Table 3.1. Standard errors in parentheses are clustered by date of letter receipt. p-values adjusted for multiple hypothesis testing (List et al., 2016).

Table C7: Treatment Effects on Other Outcomes

Panel A: Tax Payment	TPR 2014	TPR 2015	TP
	% Liability Paid before Enforcement (1)	% Liability Paid before Enforcement (2)	% Liability Paid before Deadline (3)
Simplified (S)	0.003	0.011	0.002
	(0.002)	(0.004)	(0.001)
+ Deterrence	0.002	0.002	-0.000
	(0.001)	(0.002)	(0.001)
+ Tax Morale	0.000	-0.001	-0.000
	(0.002)	(0.002)	(0.001)
+ Deterrence + Tax Morale		0.003	
		(0.002)	
P-values of tests:			
Simplified=Control	0.124	0.120	0.001
S+Deterrence=Simplified	0.318	0.987	0.827
S+Tax Morale=Simplified	0.416	0.472	0.708
S+Deterrence+Tax Morale=S+Deterrence		0.715	
Control mean	0.915	0.900	0.941
N	124,032	105,934	892,310
Panel B: Tax Filing	Log pre-check total tax due (1)	Log self-employed profits (2)	Log self-employed expenses (3)
Tax Morale	-0.003	0.010	-0.014
	(0.003)	(0.016)	(0.014)
P-values of test:			
Tax Morale=Control	0.584	0.750	0.776
Control mean	13.446	12.767	12.940
N	850,778	64,606	44,919
Panel B (continued)	Log salaried expenses (4)	Log general expenses (5)	
Tax Morale	-0.004	-0.006	
	(0.006)	(0.005)	
P-values of test:			
Tax Morale=Control	0.844	0.526	
Control mean	13.155	11.082	
N	39,176	290,551	

Note: The table presents treatment effect estimates for other outcomes of interest in the tax payment (TP FY2016 Panel A) and the tax filing (TF FY2016 Panel B) experiments. In Panel A the sample consists of late payers who had made some payment before enforcement started. Control variables are listed in Table 3.1. Robust standard errors in parentheses, clustered by date of letter receipt in Panel A. p-values adjusted for multiple hypothesis testing (List et al., 2016).

Table C8: Tax Filing: Survey Results

	Knows how taxes are spent	Knowledge index	Agrees with how taxes are spent	Agreement index	Values public services	Satisfied with tax system	Agrees should be honest
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tax Morale	0.114	0.101	0.032	0.025	0.045	0.006	0.009
	(0.008)	(0.010)	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)
P-values of test:							
Tax Morale=Control	0.000	0.000	0.000	0.025	0.000	0.442	0.382
N	66,530	47,194	66,430	47,897	66,698	66,874	66,607

Note: The table presents treatment effect estimates from the analysis of survey responses in the tax filing experiment (TF FY2016). Outcomes are standardized using mean and standard deviation of control group responses. Control variables are dummies for gender and age categories. Robust standard errors in parentheses. p-values adjusted for multiple hypothesis testing (List et al., 2016).

Table C9: Heterogeneous Effects – Payment Reminder Experiment FY2014

Probability of payment before enforcement			
Simplified (S)	0.041		
	(0.040)		
+ Deterrence	0.080		(continued)
	(0.030)		
+ Tax Morale	0.090	Solvency score Q3 * Simplified	0.056
	(0.025)		(0.014)
Age 31-40 * Simplified	0.012	* S + Deterrence	-0.004
	(0.013)		(0.008)
* S + Deterrence	-0.011	* S + Tax Morale	-0.004
	(0.012)		(0.011)
* S + Tax Morale	-0.006	Solvency score Q4 * Simplified	0.024
	(0.010)		(0.016)
Age 41-50 * Simplified	0.025	* S + Deterrence	-0.016
	(0.014)		(0.013)
* S + Deterrence	-0.026	* S + Tax Morale	-0.001
	(0.012)		(0.015)
* S + Tax Morale	-0.026	Solvency score Q5 * Simplified	-0.030
	(0.011)		(0.021)
Age 51-60 * Simplified	0.011	* S + Deterrence	-0.002
	(0.013)		(0.011)
* S + Deterrence	-0.028	* S + Tax Morale	0.012
	(0.010)		(0.011)
* S + Tax Morale	-0.028	Liability Q2 * Simplified	-0.050
	(0.010)		(0.012)
Age 61+ * Simplified	-0.017	* S + Deterrence	-0.005
	(0.013)		(0.010)
* S + Deterrence	-0.024	* S + Tax Morale	0.017
	(0.013)		(0.011)
* S + Tax Morale	-0.016	Liability Q3 * Simplified	-0.042
	(0.008)		(0.010)
One child * Simplified	0.019	* S + Deterrence	-0.019
	(0.013)		(0.007)
* S + Deterrence	0.008	* S + Tax Morale	0.004
	(0.010)		(0.007)
* S + Tax Morale	0.013	Liability Q4 * Simplified	-0.062
	(0.012)		(0.010)
Two or more children * Simplified	0.027	* S + Deterrence	-0.016
	(0.014)		(0.010)
* S + Deterrence	-0.011	* S + Tax Morale	0.016
	(0.012)		(0.010)
* S + Tax Morale	-0.012	Liability Q5 * Simplified	-0.046
	(0.011)		(0.011)
Solvency score Q2 * Simplified	0.059	* S + Deterrence	-0.041
	(0.011)		(0.008)
* S + Deterrence	-0.008	* S + Tax Morale	0.007
	(0.007)		(0.010)
* S + Tax Morale	-0.013		
	(0.005)	N	229,751

Note: The table presents the heterogeneous treatment effects of the TPR FY2015 experiment. Control variables are listed in Table 3.1. The full set of interactions between individual control and treatment variables are included in the estimation (coefficients not reported). Standard errors in parentheses are clustered by date of letter receipt.

Table C10: Heterogeneous Effects – Payment Reminders Experiment FY2015

Probability of payment before enforcement			
Simplified (S)	0.079		
	(0.066)		
+ Deterrence	0.042		(continued)
	(0.047)		
+ Tax Morale	-0.002	Solvency score Q3 * Simplified	-0.018
	(0.058)		(0.013)
Age 31-40 * Simplified	-0.017	* S + Deterrence	-0.021
	(0.013)		(0.013)
* S + Deterrence	0.031	* S + Tax Morale	-0.027
	(0.016)		(0.013)
* S + Tax Morale	0.008	Solvency score Q4 * Simplified	-0.002
	(0.012)		(0.021)
Age 41-50 * Simplified	-0.002	* S + Deterrence	0.001
	(0.013)		(0.010)
* S + Deterrence	-0.015	* S + Tax Morale	0.016
	(0.009)		(0.012)
* S + Tax Morale	0.008	Solvency score Q5 * Simplified	0.109
	(0.013)		(0.011)
Age 51-60 * Simplified	-0.011	* S + Deterrence	0.055
	(0.012)		(0.024)
* S + Deterrence	-0.012	* S + Tax Morale	-0.001
	(0.010)		(0.017)
* S + Tax Morale	0.005	Liability Q2 * Simplified	0.037
	(0.009)		(0.015)
Age 61+ * Simplified	-0.004	* S + Deterrence	0.001
	(0.010)		(0.017)
* S + Deterrence	-0.007	* S + Tax Morale	-0.015
	(0.015)		(0.012)
* S + Tax Morale	-0.006	Liability Q3 * Simplified	-0.021
	(0.013)		(0.008)
One child * Simplified	-0.015	* S + Deterrence	0.047
	(0.020)		(0.013)
* S + Deterrence	-0.007	* S + Tax Morale	0.009
	(0.017)		(0.015)
* S + Tax Morale	-0.014	Liability Q4 * Simplified	-0.009
	(0.019)		(0.009)
Two or more children * Simplified	0.070	* S + Deterrence	-0.009
	(0.010)		(0.009)
* S + Deterrence	-0.033	* S + Tax Morale	0.088
	(0.013)		(0.016)
* S + Tax Morale	-0.008	Liability Q5 * Simplified	0.012
	(0.015)		(0.013)
Solvency score Q2 * Simplified	0.020	* S + Deterrence	0.005
	(0.014)		(0.012)
* S + Deterrence	-0.001	* S + Tax Morale	0.009
	(0.011)		(0.012)
* S + Tax Morale	-0.026		
	(0.010)	N	0.089

Note: The table presents the heterogeneous treatment effects of the TPR FY2015 experiment. Control variables are listed in Table 3.1. The full set of interactions between individual control and treatment variables are included in the estimation (coefficients not reported). Estimates for Deterrence and Tax Morale joint treatment omitted for brevity. Standard errors in parentheses are clustered by date of letter receipt.

Table C11: Number of Follow-up Enforcements FY2014

	Nr registered letters within 180 days (1)	Nr garnishments within 180 days (2)	Nr bailiffs within 180 days (3)
Simplified	-0.074	-0.028	-0.012
	(0.003)	(0.002)	(0.002)
Control mean	0.350	0.134	0.078
N	229,751	229,751	229,751

Note: The table presents treatment effect estimates on the number of enforcement actions by type from the payment reminders experiment (TPR FY2014). Control variables are listed in Table 3.1. Standard errors in parentheses are clustered by date of letter receipt.

Table C12: RDD: Number of Follow-up Enforcements FY2014

	Nr registered letters within 180 days (1)	Nr garnishments within 180 days (2)	Nr bailiffs within 180 days (3)
Simplified	-0.070	-0.019	0.002
	(0.018)	(0.015)	(0.003)
Enforcement	0.110	0.071	0.000
	(0.025)	(0.021)	(0.004)
Simplified*Enforcement	-0.057	-0.032	-0.000
	(0.027)	(0.022)	(0.005)
Control mean	0.159	0.061	0.002
N	25,855	20,338	30,348

Note: The table presents treatment effect estimates from the regression discontinuity design analysis embedded in the payment reminder experiment (TPR FY2014). Simplified is a dummy variable equal to one for taxpayers who received a simplified letter. Enforcement is a dummy variable equal to one for liability amounts above the cut-off value. Control variables are listed in Table 3.1. Standard errors in parentheses are clustered by date of letter receipt.

Table C13: Replication of TPR experiment in FY2015

	Probability of some payment at 14 days (before enforcement)
Simplified (S)	0.107
	(0.004)
+ Deterrence	0.014
	(0.003)
+ Tax Morale	-0.012
	(0.003)
+ Deterrence & Tax Morale	0.006
	(0.003)
P-values of tests:	
Simplified=Control	0.001
S + Deterrence=Simplified	0.001
S + Tax Morale=Simplified	0.007
S + Deterrence + Tax Morale= S + Deterrence	0.011
Control mean	0.418
N	202,730

Note: The table presents results from the FY2015 TPR experiment, which replicated the FY2014 TPR experiment. The sample is the universe of late payers in FY2015. Control variables are listed in Table 3.1. Standard errors in parentheses are clustered by date of letter receipt. p-values adjusted for multiple hypothesis testing (List et al., 2016).

Table C14: Repeated Treatment Effects

	Probability of some payment 14 days / follow-up	
	Sample of Taxpayers late in FY2014 (2)	Sample of Taxpayers late in FY2014 and FY2015 (1)
Simplified 2014 (S 2014)	0.001	0.011
	(0.008)	(0.012)
+ Deterrence 2014 (D 2014)	-0.002	-0.010
	(0.007)	(0.010)
+ Tax Morale 2014 (TM 2014)	-0.003	-0.016
	(0.008)	(0.018)
Simplified 2015 (S 2015)	0.025	0.107
	(0.009)	(0.013)
+ Deterrence 2015 (D 2015)	0.001	0.005
	(0.004)	(0.012)
+ Tax Morale 2015 (TM 2015)	-0.004	-0.024
	(0.004)	(0.007)
S 2014 * S 2015	0.005	-0.020
	(0.010)	(0.015)
S 2014 * S + D 2015	-0.002	-0.011
	(0.005)	(0.017)
S 2014 * S + TM 2015	0.004	0.025
	(0.004)	(0.015)
S + D 2014 * S 2015	0.000	0.016
	(0.009)	(0.014)
S + D 2014 * S + D 2015	0.002	0.010
	(0.005)	(0.011)
S + D 2014 * S + TM 2015	-0.002	-0.021
	(0.003)	(0.014)
S + TM 2014 * S 2015	-0.007	0.005
	(0.009)	(0.021)
S + TM 2014 * S + D 2015	0.007	0.027
	(0.005)	(0.011)
S + TM 2014 * S + TM 2015	-0.002	-0.010
	(0.002)	(0.007)
P-values of tests:		
S 2014 = Control	0.824	0.971
S 2015 = Control	0.308	0.001
S 2014 * S 2015 = S 2015	0.823	0.992
S + D 2014 = S 2014	0.947	0.945
S + D 2015 = S 2015	0.907	0.996
S + D 2014 * S + D 2015 = S 2014 * S + D 2015	0.986	0.582
S + TM 2014 = S 2014	0.976	0.751
S + TM 2015 = S 2015	0.993	0.948
S + TM 2014 * S + TM 2015 = S 2014 * S + TM 2015	0.165	0.959
Control mean	0.825	0.410
N	229,751	66,705

Note: The table present treatment effect estimates for repeated treatment in the payment reminders experiment. Sample size is limited to individuals who were late with payment in both FY2014 and FY2015. For FY2015 treatment assignment both dummies for Deterrence and Tax Morale equal one for individuals who received a letter with both a deterrence and tax morale message. Control variables are listed in Table 3.1. Standard errors in parentheses are clustered by date of letter receipt.

References

Abeler, J. and S. Jäger (2015). Complex tax incentives. *American Economic Journal: Economic Policy 7*(3), 1–28.

Aghion, P., U. Akcigit, M. Lequien, and S. Stantcheva (2017). Tax simplicity and heterogeneous learning. Technical report, National Bureau of Economic Research.

Alcott, H., B. Lockwood, and D. Taubinsky (2018). Regressive sin taxes, with an application to the optimal soda tax. Working paper.

Alesina, A., A. Miano, and S. Stantcheva (2018). Immigration and redistribution. NBER Working Papers 24733, National Bureau of Economic Research, Inc.

Allcott, H. and T. Rogers (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review 104*(10), 3003–37.

Allingham, M. and A. Sandmo (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics 1*, 323–338.

Allport, G. W., K. Clark, and T. Pettigrew (1954). *The nature of prejudice*. Addison-Wesley Reading, MA.

Alm, J. (2012). Measuring, explaining, and controlling tax evasion: Lessons from theory, experiments, and field studies. *International Tax and Public Finance 19*(1), 54–77.

Alstadsæter, A., N. Johannesen, and G. Zucman (2018). Who owns the wealth in tax havens? macro evidence and implications for global inequality. *Journal of Public Economics 162*, 89–100.

Altonji, J. and D. Card (1991). The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, Trade, and the Labor Market*, pp. 201–234. National Bureau of Economic Research, Inc.

Altonji, J. G., P. Arcidiacono, and A. Maurel (2016). The analysis of field choice in college and graduate school: Determinants and wage effects. In *Handbook of the Economics of Education*, Volume 5, pp. 305–396. Elsevier.

Andreoni, J., B. Erard, and J. Feinstein (1998). Tax compliance. *Journal of Economic Literature 36*, 818–860.

Anelli, M., K. Shih, and K. Williams (2018). Foreign Peer Effects and STEM Major Choice. Working paper.

Angrist, J. D. and A. D. Kugler (2003). Protective or counter-productive? Labour market institutions and the effect of immigration on natives. *The Economic Journal 113*(488), F302–F331.

Arrow, K. J. (1973). Higher education as a filter. *Journal of Public Economics 2*(3), 193 – 216.

Ayllon, S. and N. Nollenberger (2016). Are recessions good for human capital accumulation? NEGOTIATE working paper no. 5.1.

Barone, G., A. D'Ignazio, G. de Blasio, and P. Naticchioni (2016). Mr. Rossi, Mr. Hu and politics. The role of immigration in shaping natives' voting behavior. *Journal of Public Economics 136*(C), 1–13.

Bartik, T. (1991). *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute for Employment Research.

Basten, C. and M. Siegenthaler (2019). Do immigrants take or create residents jobs? Evidence from Free Movement of Workers in Switzerland. *Scandinavian Journal of Economics 121*(3), 994–1019.

Becker, G. (1968). Crime and punishment: An economic approach. *Journal of Political Economy 76*.

Beerli, A., J. Ruffner, M. Siegenthaler, and G. Peri (2018, November). The Abolition of Immigration Restrictions and the Performance of Firms and Workers: Evidence from Switzerland. Working Paper 25302, National Bureau of Economic Research.

Beffy, M., D. Fougre, and A. Maurel (2012, February). Choosing the Field of Study in Postsecondary Education: Do Expected Earnings Matter? *The Review of Economics and Statistics 94*(1), 334–347.

Benartzi, S., J. Beshears, K. L. Milkman, C. R. Sunstein, R. H. Thaler, M. Shankar, W. Tucker-Ray, W. J. Congdon, and S. Galing (2017). Should governments invest more in nudging? *Psychological science 28*(8), 1041–1055.

Benzarti, Y. (2017). How taxing is tax filing? using revealed preferences to estimate compliance costs. Working Paper 23903, National Bureau of Economic Research.

Besley, T., A. Jensen, and T. Persson (2019). Norms, enforcement, and tax evasion. Working Paper 25575, National Bureau of Economic Research.

Betts, J. (1998). Educational crowding out: Do immigrants affect the educational attainment of american minorities? University of California at San Diego, Economics Working Paper Series, Department of Economics, UC San Diego.

Bhargava, S. and D. Manoli (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an irs field experiment. *American Economic Review 105*(11), 3489–3529.

Blumenthal, M., C. Christian, J. Slemrod, and M. G. Smith (2001). Do normative appeals affect tax compliance? evidence from a controlled experiment in minnesota. *National Tax Journal*, 125–138.

Boeri, T. and H. Brücker (2005). Why are Europeans so tough on migrants? *Economic Policy 20*(44), 629–703.

Borjas, G. J. (2003, 11). The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market. *The Quarterly Journal of Economics 118*(4), 1335–1374.

Borjas, G. J. (2004, March). Do foreign students crowd out native students from graduate programs? Working Paper 10349, National Bureau of Economic Research.

Borusyak, K., P. Hull, and X. Jaravel (2018). Quasi-experimental shift-share research designs. Papers, arXiv.org.

Bott, K. M., A. W. Cappelen, E. Sorensen, and B. Tungodden (2017). You've got mail: a randomised field experiment on tax evasion. Department of Economics, Norwegian School of Economics Discussion Paper 26/2014.

Brockmeyer, A., M. Hernandez, S. Kettle, and S. Smith (2019). Casting a wider tax net: Experimental evidence from costa rica. *American Economic Journal: Economic Policy*.

Brücker, H., A. Hauptmann, E. Jahn, and R. Upward (2014). Migration and imperfect labor markets: Theory and cross-country evidence from Denmark, Germany and the UK. *European Economic Review 66*(C), 205–225.

Brücker, H. and E. Jahn (2011). Migration and Wage-setting: Reassessing the Labor Market Effects of Migration. *Scandinavian Journal of Economics 113*, 286–317.

Brunner, B. and A. Kuhn (2018). Immigration, cultural distance and natives' attitudes towards immigrants: Evidence from Swiss voting results. *Kyklos 71*(1), 28–58.

Card, D. (1996). The effect of unions on the structure of wages: A longitudinal analysis. *Econometrica 64*(4), 957–79.

Card, D. (2001). Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics 19*(1), 22–64.

Card, D. and S. De La Rica (2006). Firm-Level contracting and the structure of wages in Spain. *ILR Review 59*(4), 573–592.

Card, D., C. Dustmann, and I. Preston (2012). Immigration, wages, and compositional amenities. *Journal of the European Economic Association 10*(1), 78–119.

Cavaille, C. and J. Ferwerda (2017). How distributional conflict over public spending drives support for anti-immigrant parties. Cage online working paper series, Competitive Advantage in the Global Economy (CAGE).

Cavaille, C. and J. Marshall (2019). Education and anti-immigration attitudes: Evidence from compulsory schooling reforms across western europe. *American Political Science Review 113*(01), 254–263.

Chernozhukov, V., M. Demirer, E. Duflo, and I. Fernandez-Val (2018). Generic machine learning inference on heterogenous treatment effects in randomized experiments. NBER Working Paper 24678, National Bureau of Economic Research.

Chetty, R. and E. Saez (2013). Teaching the tax code: Earnings responses to an experiment with eitc recipients. *American Economic Journal: Applied Economics 5*(1), 1–31.

Cortés, P. and J. Pan (2015). Immigration and Occupational Choice of Natives: the Case of Nurses in the United States. *CESifo Economic Studies 61*(3-4), 797–823.

Cox, J. C., D. Kreisman, and S. Dynarski (2018). Designed to fail: Effects of the default option and information complexity on student loan repayment. Working Paper 25258, National Bureau of Economic Research.

Cranor, T., J. Goldin, T. Homonoff, and L. Moore (2018). Communicating tax penalties to delinquent taxpayers: Evidence from a field experiment. Working paper.

Cronqvist, H., R. H. Thaler, and F. Yu (2018). When nudges are forever: Inertia in the swedish premium pension plan. In *AEA Papers and Proceedings*, Volume 108, pp. 153–58.

D'Amuri, F. and G. Peri (2014). Immigration, jobs, and employment protection: Evidence from Europe before and during the Great Recession. *Journal of the European Economic Association 12*(2), 432–464.

Del Carpio, L. (2014). Are the neighbors cheating? evidence from a social norm experiment on property taxes in peru. Economics Department, Princeton University.

Dellas, H. and P. Sakellaris (2003). On the cyclicality of schooling: Theory and evidence. *Oxford Economic Papers 55*(1), 148–172.

Denzler, S. and S. C. Wolter (2011). Too Far to Go? Does Distance Determine Study Choices? Working Paper 5712, IZA.

d'Hombres, B. and L. Nunziata (2016). Wish you were here? Quasi-experimental evidence on the effect of education on self-reported attitude toward immigrants. *European Economic Review 90*(C), 201–224.

Di Tella, R., R. Perez-Truglia, A. Babino, and M. Sigman (2015). Conveniently upset: Avoiding altruism by distorting beliefs about others' altruism. *American Economic Review 105*(11), 3416–42.

Diehl, C., T. Hinz, and K. Auspurg (2018). Who is afraid of skilled migrants from Europe? Exploring support for immigration control in Switzerland. *Swiss Journal of Sociology 44*(1).

Dustmann, C., T. Frattini, and I. Preston (2013). The effect of immigration along the distribution of wages. *Review of Economic Studies 80*(1), 145–173.

Dustmann, C. and I. Preston (2001). Attitudes to ethic minorities, ethnic context and location decisions. *Economic Journal 111*(470), 353–73.

Dustmann, C. and I. Preston (2007). Racial and economic factors in attitudes to immigration. *The B.E. Journal of Economic Analysis & Policy 7*(1), 1–41.

Dustmann, C., U. Schönberg, and J. Stuhler (2016a). The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives 30*(4), 31–56.

Dustmann, C., U. Schönberg, and J. Stuhler (2016b). The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives 30*(4), 31–56.

Dustmann, C., U. Schönberg, and J. Stuhler (2017). Labor supply shocks, native wages, and the adjustment of local employment. *The Quarterly Journal of Economics 132*(1), 435–483.

Dustmann, C., U. Schnberg, and J. Stuhler (2017). Labor supply shocks, native wages, and the adjustment of local employment. *The Quarterly Journal of Economics 132*(1), 435–483.

Dustmann, C., K. Vasiljeva, and A. Damm (2019). Refugee migration and electoral outcomes. *The Review of Economic Studies*.

Dwenger, N., H. Kleven, I. Rasul, and J. Rincke (2016, August). Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany. *American Economic Journal: Economic Policy 8*(3), 203–232.

Edo, A. (2016). How do rigid labor markets absorb immigration? Evidence from France. *IZA Journal of Migration and Development 5*(1), 1–20.

Edo, A., Y. Giesing, J. Öztunc, and P. Poutvaara (2019). Immigration and electoral support for the far-left and the far-right. *European Economic Review 115*, 99–143.

Edo, A. and H. Rapoport (2019). Minimum wages and the labor market effects of immigration. *Labour Economics 61*, 101753.

Facchini, G. and A. M. Mayda (2009). Does the welfare state affect individual attitudes toward immigrants? Evidence across countries. *The Review of Economics and Statistics 91*(2), 295–314.

Facchini, G. and M. Steinhardt (2011). What drives u.s. immigration policy? evidence from congressional roll call votes. *Journal of Public Economics 95*(7-8), 734–743.

Fellner, G., R. Sausgruber, and C. Traxler (2013). Testing enforcement strategies in the field: Threat, moral appeal and social information. *Journal of the European Economic Association 11*(3), 634–660.

Foged, M., L. Hasager, and V. Yasenov (2019). The role of institutions in the labor market impact of immigration.

Gerfin, M. and B. Kaiser (2010, Oct). The effects of immigration on wages: An application of the structural skill-cell approach. *Swiss Journal of Economics and Statistics 146*(4), 709–739.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (forthcoming). Bartik instruments: What, when, why, and how. *American Economic Review*.

Guiso, L., H. Herrera, M. Morelli, and T. Sonno (2017). Populism: Demand and supply. *CEPR Discussion Paper No. DP11871*.

Gürtzgen, N. (2016). Estimating the wage premium of collective wage contracts: Evidence from longitudinal linked employer–employee data. *Industrial Relations: A Journal of Economy and Society 55*(2), 294–322.

Guyton, J., D. S. Manoli, B. Schafer, and M. Sebastiani (2016). Reminders & recidivism: evidence from tax filing & eitc participation among low-income nonfilers. Working Paper 21904, National Bureau of Economic Research.

Hainmueller, J. and D. Hangartner (2013). Who gets a Swiss passport? A natural experiment in immigrant discrimination. *American Political Science Review 107*(01), 159–187.

Hainmueller, J. and D. J. Hopkins (2014). Public attitudes toward immigration. *Annual Review of Political Science 17*, 225–249.

Halla, M., A. F. Wagner, and J. Zweimüller (2017). Immigration and voting for the far right. *Journal of the European Economic Association 15*(6), 1341–1385.

Hallsworth, M., J. A. List, R. D. Metcalfe, and I. Vlaev (2015). The making of homo honoratus: from omission to commission. Working Paper 21210, National Bureau of Economic Research.

Hallsworth, M., J. A. List, R. D. Metcalfe, and I. Vlaev (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics 148*, 14–31.

Hanson, G. H. and C. Liu (2016). High-Skilled Immigration and the Comparative Advantage of Foreign Born Workers across US Occupations. Working paper.

Hanson, G. H. and M. J. Slaughter (2016). High-skilled immigration and the rise of STEM occupations in U.S. employment. Working Paper 22623, National Bureau of Economic Research.

Hanspeter Kriesi, Matthias Brunner, Franois Lortan (Vox Survey) (2019). Standardized Surveys VoxIt 1981-2016 [Dataset]. Distributed by FORS, Lausanne.

HM Revenue & Customs (2018). Measuring tax gaps 2018 edition. Technical report, HM Revenue & Customs.

Hoopes, J. L., D. H. Reck, and J. Slemrod (2015). Taxpayer search for information: Implications for rational attention. *American Economic Journal: Economic Policy 7*(3), 177–208.

Hoxby, C. M. (2009, December). The Changing Selectivity of American Colleges. *Journal of Economic Perspectives 23*(4), 95–118.

Hunt, J. (2017). The impact of immigration on the educational attainment of natives. *Journal of Human Resources 52*(4), 1060–1118.

Hunt, J. and M. Gauthier-Loiselle (2010, April). How Much Does Immigration Boost Innovation? *American Economic Journal: Macroeconomics 2*(2), 31–56.

Internal Revenue Service (2016). Tax gap estimates for tax years 2008–2010. Technical report, Internal Revenue Service.

Jackson, O. (2018). Does immigration crowd natives into or out of higher education? Working Papers 15-18, Federal Reserve Bank of Boston.

Jaeger, D., J. Ruist, and J. Stuhler (2018). Shift-share instruments and the impact of immigration. NBER Working Papers 24285, National Bureau of Economic Research, Inc.

Jensen, A. (2019). Employment structure and the rise of the modern tax system. Working Paper 25502, National Bureau of Economic Research.

Jensen, R. (2010, 05). The (Perceived) Returns to Education and the Demand for Schooling. *The Quarterly Journal of Economics 125*(2), 515–548.

John, P. and T. Blume (2018). How best to nudge taxpayers? the impact of message simplification and descriptive social norms on payment rates in a central london local authority. *Journal of Behavioral Public Administration 1*(1).

Johnson, M. T. (2013). The impact of business cycle fluctuations on graduate school enrollment. *Economics of Education Review 34*(C), 122–134.

Kahn, L. M. (2000). Wage inequality, collective bargaining, and relative employment from 1985 to 1994: Evidence from fifteen OECD countries. *Review of Economics and Statistics 82*(4), 564–579.

Keen, M. and J. Slemrod (2017). Optimal tax administration. *Journal of Public Economics 152*, 133–142.

Kirchler, E. (2007). *The economic psychology of tax behaviour*. New York, NY: Cambridge University Press.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or unable to cheat? evidence from a tax audit experiment in denmark. *Econometrica 79*(3), 651–692.

Kleven, H. J. and W. Kopczuk (2011). Transfer program complexity and the take-up of social benefits. *American Economic Journal: Economic Policy 3*(1), 54–90.

Kleven, H. J., C. T. Kreiner, and E. Saez (2016). Why can modern governments tax so much? an agency model of firms as fiscal intermediaries. *Economica 83*(330), 219–246.

Lemieux, T. (1998). Estimating the effects of unions on wage inequality in a panel data model with comparative advantage and nonrandom selection. *Journal of Labor Economics 16*(2), 261–91.

List, J. A., A. M. Shaikh, and Y. Xu (2016, January). Multiple hypothesis testing in experimental economics. Working Paper 21875, National Bureau of Economic Research.

Llull, J. (2018, 09). Immigration, Wages, and Education: A Labour Market Equilibrium Structural Model. *The Review of Economic Studies 85*(3), 1852–1896.

Long, B. T. (2014, January). *The Financial Crisis and College Enrollment: How Have Students and Their Families Responded?*, pp. 209–233. University of Chicago Press.

Lucas, R. E. J. (1988). On the mechanics of development. *Journal of Monetary Economics 22*, 3–24.

Luttmer, E. F. and M. Singhal (2014). Tax morale. *Journal of Economic Perspectives 28*(4), 149–168.

Machin, S. and R. Murphy (2017, 04). Paying out and crowding out? The globalization of higher education. *Journal of Economic Geography 17*(5), 1075–1110.

Magruder, J. (2012). High unemployment yet few small firms: The role of centralized bargaining in South Africa. *American Economic Journal: Applied Economics 4*(3), 138–66.

Margaryan, S., A. Paul, and T. Siedler (forthcoming). Does Education Affect Attitudes Towards Immigration? Evidence from Germany. *Journal of Human Resources*.

Mayda, A. M. (2006). Who is against immigration? A cross-country investigation of individual attitudes toward immigrants. *The Review of Economics and Statistics 88*(3), 510–530.

Mayda, A. M., G. Peri, and W. Steingress (2018). The political impact of immigration: Evidence from the United States. NBER Working Papers 24510, National Bureau of Economic Research, Inc.

McHenry, P. (2015). Immigration and the human capital of natives. *Journal of Human Resources 50*(1), 34–71.

Naritomi, J. (2018). Consumers as tax auditors. *CEPR Discussion Paper No. DP13276*.

OECD (2010). Understanding and influencing taxpayers' compliance behavior. Technical report, OECD Centre for Tax Policy and Administration.

OECD (2020). Foreign population (indicator). accessed 01.03.2020.

O'Rourke, K. and R. Sinnott (2006). The determinants of individual attitudes towards immigration. *European Journal of Political Economy 22*(4), 838–861.

Orrenius, P. M. and M. Zavodny (2015). Does immigration affect whether US natives major in science and engineering? *Journal of Labor Economics 33*(S1), S79–S108.

Ortega, F. and J. G. Polavieja (2012). Labor-market exposure as a determinant of attitudes toward immigration. *Labour Economics 19*(3), 298–311.

Ottaviano, G. and G. Peri (2012a). Rethinking the effect of immigration on wages. *Journal of the European Economic Association 10*(1), 152–197.

Ottaviano, G. and G. Peri (2012b). Rethinking the effect of immigration on wages. *Journal of the European Economic Association 10*(1), 152–197.

Otto, A. and M. Steinhardt (2014). Immigration and election outcomes Evidence from city districts in Hamburg. *Regional Science and Urban Economics 45*(C), 67–79.

Pecoraro, M. and D. Ruedin (2019). Occupational exposure to foreigners and attitudes towards equal opportunities. *Migration Studies*.

Perez-Truglia, R. and U. Troiano (2018). Shaming tax delinquents. *Journal of Public Economics 167*, 120 – 137.

Peri, G. and C. Sparber (2011). Highly Educated Immigrants and Native Occupational Choice. *Industrial Relations: A Journal of Economy and Society 50*(3), 385–411.

Pomeranz, D. (2015). No taxation without information: Deterrence and self-enforcement in the value added tax. *American Economic Review 105*(8), 2539–69.

Ransom, T. and J. V. Winters (2020). Do Foreigners Crowd Natives out of STEM Degrees and Occupations? Evidence from the US Immigration Act of 1990. *ILR Review*.

Rodrik, D. (1997). *Has Globalization Gone Too Far?* Washington, DC: Institute for International Economics.

Scheve, K. F. and M. J. Slaughter (2001). Labor market competition and individual preferences over immigration policy. *The Review of Economics and Statistics 83*(1), 133–145.

Schindler, D. and M. Westcott (2017). Shocking racial attitudes: Black g.i.s in europe. CESifo Working Paper Series 6723, CESifo Group Munich.

Schündeln, M. (2014). Are Immigrants More Mobile than Natives? Evidence from Germany. *Journal of Regional Science 54*(1), 70–95.

Schweri, J. and J. Hartog (2017). Do wage expectations predict college enrollment? Evidence from healthcare. *Journal of Economic Behavior & Organization 141*, 135 – 150.

Sciarini, P., F. Cappelletti, A. C. Goldberg, and S. Lanz (2016). The underexplored species: Selective participation in direct democratic votes. *Swiss Political Science Review 22*(1), 75–94.

Shih, K. (2017). Do international students crowd-out or cross-subsidize Americans in higher education? *Journal of Public Economics 156*, 170 – 184.

Slemrod, J. (2018). Tax compliance and enforcement. Working Paper 24799, National Bureau of Economic Research.

Slemrod, J., M. Blumenthal, and C. Christian (2001). Taxpayer response to an increased probability of audit: evidence from a controlled experiment in minnesota. *Journal of public economics 79*(3), 455–483.

Slemrod, J. and S. Yitzhaki (2002). Tax avoidance, evasion, and administration. In A. J. Auerbach and M. Feldstein (Eds.), *Handbook of Public Economics*, Volume 3 of *Handbook of Public Economics*, Chapter 22, pp. 1423–1470. Elsevier.

Steinmayr, A. (forthcoming). Contact versus exposure: Refugee presence and voting for the far-right. *Review of Economics and Statistics*.

Strubi, P., J. Vesela, and J. Babel (2018). *Übergänge nach Abschluss der Sekundarstufe II und Integration in den Arbeitsmarkt*. Bundesamt für Statistik.

Tabellini, M. (2019, 05). Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration. *The Review of Economic Studies 87*(1), 454–486.

Torgler, B. (2007). *Tax compliance and tax morale: A theoretical and empirical analysis*. Northampton, MA: Edward Elgar Publishing.

United Nations, Department of Economic and Social Affairs, Population Division (UNPD) (2019). International Migrant Stock 2019 (United Nations database, POP/DB/MIG/Stock/Rev.2019).

Wager, S. and S. Athey (2018). Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association 113*(523), 1228–1242.

Wiswall, M. and B. Zafar (2015). Determinants of College Major Choice: Identification using an Information Experiment. *The Review of Economic Studies 82*(2), 791–824.

Xia, X. (2016). Forming wage expectations through learning: Evidence from college major choices. *Journal of Economic Behavior & Organization 132*(PA), 176–196.